

FOR OFFICIAL USE ONLY

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

20548
SEARCH REQUEST FORM

Examiner # (Mandatory): 76554 Requester's Full Name: Li, Leo

Art Unit 1645 Location (Bldg/Room#): 8E17 Phone (circle 305-306-308) 8891

Serial Number: 09/107,979 Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Date: _____

Keywords (include any known synonyms registry numbers, explanation of initialisms):

please search

SEQ ID NO = 4

TECH/CHEM. DIVISION
(STIC)

OCT 12 1999

RECEIVED

Search Topic:

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

please Rush it

AC
ANTHONY C. CAPUTA
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 1600

FOR OFFICIAL USE ONLY

STAFF USE ONLY

Searcher: <u>Wang</u>	Type of Search	Vendors (include cost where applicable)
Searcher Phone #: _____	<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> STN
Searcher Location: _____	<input checked="" type="checkbox"/> A.A. Sequence	<input type="checkbox"/> Questel/Orbit
Date Picked Up: <u>10/13</u>	<input type="checkbox"/> Structure (#)	<input type="checkbox"/> Lexis/Nexis
Date Completed: _____	<input type="checkbox"/> Bibliographic	<input type="checkbox"/> WWW/Internet
Clerical Prep Time: _____	<input type="checkbox"/> Litigation 1	<input type="checkbox"/> In-house sequence systems (list)
Terminal Time: <u>21</u>	<input type="checkbox"/> Fulltext	<input type="checkbox"/> Dialog
Number of Databases: _____	<input type="checkbox"/> Procurement	<input type="checkbox"/> Dr. Link
	<input type="checkbox"/> Other	<input type="checkbox"/> Westlaw
		<input checked="" type="checkbox"/> Other (specify) <u>CompuGen</u>

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 1999, 22:41:30 ; Search time 14.23 Seconds
(without alignments)
78.232 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFKPCRDLAYCLNDGECF.....SHKHCRCKEGYGVRCDOFL 47

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	277	100.0	713	1 W97617	Mouse neuregulin I
2	277	100.0	362	1 W97620	Mouse neuregulin I
3	277	100.0	720	1 W97618	Human neuregulin I
4	277	100.0	360	1 W97621	Human neuregulin I
5	277	100.0	47	1 W97622	Human neuregulin I
6	277	100.0	696	1 W97619	Human neuregulin I
7	277	100.0	157	1 Y05451	Human neuregulin-III
8	277	100.0	720	1 Y05452	Human neuregulin-III
9	116.5	42.1	52	1 W05182	Neu differentiation
10	113.5	41.0	263	1 R28537	GGF2BPP2.CDS prote
11	113.5	41.0	669	1 R29570	Human neuregulin-al
12	113.5	41.0	63	1 R45918	EGFL2. Glial growth
13	113.5	41.0	88	1 R45921	EGFL5. Glial growth
14	113.5	41.0	83	1 R46922	EGFL6. Glial growth
15	113.5	41.0	63	1 R55659	EGFL2. Glial mitog
16	113.5	41.0	88	1 R55662	EGFL5. Glial mitog
17	113.5	41.0	83	1 R55663	EGFL6. Glial mitog
18	113.5	41.0	263	1 R46896	GGF2BPP2. Glial gr
19	113.5	41.0	280	1 R45915	GGF2BPP2. Glial gr
20	113.5	41.0	280	1 R55656	GGF2BPP2. Glial mi
21	113.5	41.0	263	1 R55689	GGF2BPP2. Glial mi
22	113.5	41.0	125	1 R69564	Human NDF-alpha3 c
23	113.5	41.0	462	1 R69562	Human NDF-alpha2b
24	113.5	41.0	462	1 R68561	Human proNDF-alpha
25	113.5	41.0	280	1 R67244	Bovine glial cell
26	113.5	41.0	63	1 R67250	Human epidermal li
27	113.5	41.0	88	1 R67253	Human epidermal li
28	113.5	41.0	83	1 R67254	Human epidermal li
29	113.5	41.0	263	1 R67217	Putative bovine gl
30	113.5	41.0	63	1 R87461	Epidermal growth f
31	113.5	41.0	88	1 R87464	Epidermal growth f
32	113.5	41.0	83	1 R87465	Epidermal growth f
33	113.5	41.0	280	1 R87455	App2 glial growth
34	113.5	41.0	280	1 R98762	Glial growth facto
35	113.5	41.0	63	1 R96076	Epidermal growth f
36	113.5	41.0	83	1 R96080	Epidermal growth f
37	113.5	41.0	88	1 R96079	Epidermal growth f
38	113.5	41.0	375	1 W05683	Heregulin-alpha fu
39	113.5	41.0	52	1 W05184	New differentiation
40	113.5	41.0	263	1 W09360	Bovine neuregulin
41	113.5	41.0	63	1 W09363	EGFL2. Use of neur
42	113.5	41.0	88	1 W09366	EGFL5. Use of neur
43	113.5	41.0	83	1 W09367	EGFL6. Use of neur

ALIGNMENTS

RESULT	1
W97617	
ID	W97617 standard; Protein; 713 AA.
AC	W97617;
DT	10-MAY-1999 (first entry)
DE	Mouse neuregulin related ligand NRG3.
KW	Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;
KW	signal transduction; nervous system disorder; neurodegeneration;
KW	neuropathy; therapy; diagnosis.
OS	Mus sp.
FH	Key
FT	Domain
FT	Location/Qualifiers
FT	362
FT	/note= "extracellular domain, specifically claimed in Claim 5(a)"
FT	56. 91
FT	/note= "hydrophobic region"
FT	105. 286
FT	/note= "mucin-like Ser/Thr-rich region, contains sites for O-linked glycosylation"
FT	287. 334
FT	/note= "EGF-like domain"
FT	363. 385
FT	/note= "transmembrane domain"

W09902681-A1.
21-JAN-1999.
30-JUN-1998; U13411.
24-JUL-1997; US-899437.
09-JUL-1997; US-052019.
(GETH) GENENTEC INC.
Godowski PJ, Mark MR, Zhang D;
WPI; 99-120882/10
N-PSDB; X06987.
New isolated neuregulin related ligand-3 - used to develop products for treating nervous system disorders, e.g. stroke, ischaemia, infection, malignancy, Alzheimer's disease or Down's syndrome Claim 5(b); Page 59-62; 101pp; English.
This is the amino acid sequence of murine neuregulin related ligand NRG3, a novel member of the epidermal growth factor (EGF)-like family of protein ligands that binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor tyrosine phosphorylation. The sequence was deduced from the nucleotide sequences of cDNA clones (see X06987) from a mouse brain library. The EGF-like domain of NRG3 is distinct from those of NRG1 or NRG2, and NRG3 displays receptor binding characteristics that are distinct from those of other neuregulins. The invention provides human and murine NRG3 polypeptides (see also W97618), expression vectors, host cells and methods for the recombinant production of NRG3s. The NRG3 polypeptides and polynucleotides and can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, ischaemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease,

44 113.5 41.0 280 1 W09369 Human neuregulin G
45 111.5 40.3 422 1 R43651 Rat NRSE. New reco

Li lee

CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC disease, metachromatic leukodystrophy, Fabry's disease and
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC smooth muscle, such as muscular dystrophy or diseases caused by
CC skeletal or smooth muscle wasting. The products can also be used
CC for detection, diagnosis, for the production of transgenic or
CC knockout animals or for drug screening.
SQ Sequence 713 AA;

Query Match 100.0%; Score 277; DB 1; Length 713;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDFL 47
|||||
DB 288 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDFL 334
|||||

RESULT 2
W97620
ID W97620 standard; Protein: 362 AA.
AC W97620:
DT 10-MAY-1999 (first entry)
DE Mouse neuregulin related ligand NRG3 extracellular domain.
KW Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis.
OS Mus sp.
PN W09902681-A1.
PD 21-JAN-1999.
PF 30-JUN-1998; U13411.
PR 24-JUL-1997; US-899437.
PR 09-JUL-1997; US-052019.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Zhang D;
DR WPI: 99-120882/10.
PT New isolated neuregulin related ligand-3 - used to develop products
PT for treating nervous system disorders, e.g. stroke, ischaemia,
PT infection, malignancy, Alzheimer's disease or Down's syndrome
PS Claim 5(a); Page 62-63; 101pp; English.
CC This is the extracellular domain (ECD, aa1-362) of murine neuregulin
CC related ligand NRG3 (see also W97617), a novel member of the
CC epidermal growth factor (EGF)-like family of protein ligands. NRG3
CC binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,
CC activates ErbB4 receptor tyrosine phosphorylation. The invention
CC provides human and murine polypeptides (see also W97618) that have
CC at least 75% homology to the NRG3 ECD, as well as expression vectors,
CC host cells and methods for the recombinant production of novel
CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
CC enhance the survival, proliferation or differentiation of cells
CC having the ErbB4 receptor in vivo and in vitro. They can be used to
CC prevent or treat damage to a nerve or damage to other NRG3-expressing
CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC particular, they can be used to treat diseases which involve neural
CC cell growth such as demyelination, or damage or loss of glial cells
CC (e.g. multiple sclerosis). They can be used to treat patients whose
CC nervous system has been damaged by e.g. trauma, surgery, stroke,
CC ischaemia, infection, metabolic disease, nutritional deficiency,
CC malignancy, or toxic agents. NRG3 can also be used to treat
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC Gehrig's disease), Bell's palsy, conditions involving spinal
CC muscular atrophy or paralysis, neurodegenerative disorders such as
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC and Meniere's disease. They can also be used to treat neuropathies
CC associated with systemic disease including post-polio syndrome,
CC hereditary neuropathies including Charcot-Marie-Tooth disease,
CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC disease, metachromatic leukodystrophy, Fabry's disease and
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC smooth muscle, such as muscular dystrophy or diseases caused by
CC skeletal or smooth muscle wasting. The products can also be used

CC for detection, diagnosis, for the production of transgenic or
CC knockout animals or for drug screening.
SQ Sequence 362 AA;

Query Match 100.0%; Score 277; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDFL 47
|||||
DB 288 HFKPCRDLDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDFL 334
|||||

RESULT 3
W97618
ID W97618 standard; Protein: 720 AA.
AC W97618:
DT 10-MAY-1999 (first entry)
DE Human neuregulin related ligand NRG3.
KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 1..360
FT /note= "extracellular domain, specifically claimed
FT in Claim 5(a)"
FT Region 66..91
FT /note= "hydrophobic region"
FT Region 101..284
FT /note= "mucin-like Ser/Thr-rich region, contains
FT Sites for O-linked glycosylation"
FT Domain 285..354
FT /note= "EGF-like domain"
FT Domain 356..394
FT /note= "transmembrane domain"
PN W09902681-A1.
PD 21-JAN-1999. U13411.
PF 30-JUN-1998; US-899437.
PR 24-JUL-1997; US-052019.
PR 09-JUL-1997; US-052019.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Zhang D;
DR WPI: 99-120882/10.
DR N-PSDB; X06988.
PT New isolated neuregulin related ligand-3 - used to develop products
PT for treating nervous system disorders, e.g. stroke, ischaemia,
PT infection, malignancy, Alzheimer's disease or Down's syndrome
PS Claim 5(b); Page 66-69; 101pp; English.
CC This is the amino acid sequence of human neuregulin related ligand
CC NRG3, a novel member of the epidermal growth factor (EGF)-like
CC family of protein ligands that binds to the ErbB4 receptor, but not
CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor
CC tyrosine phosphorylation. The sequence was deduced from the
CC nucleotide sequence of a cDNA clone (see X06988) from a foetal brain
CC library. The EGF-like domain of NRG3 is distinct from those of NRG1
CC or NRG2, and NRG3 displays receptor binding characteristics that are
CC distinct from those of other neuregulins. An alternatively spliced
CC form of human NRG3 is provided in W97619. The invention provides
CC human and murine NRG3 polypeptides (see also W97617), expression
CC vectors, host cells and methods for the recombinant production of
CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
CC enhance the survival, proliferation or differentiation of cells
CC having the ErbB4 receptor in vivo and in vitro. They can be used to
CC prevent or treat damage to a nerve or damage to other NRG3-expressing
CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC particular, they can be used to treat diseases which involve neural
CC cell growth such as demyelination, or damage or loss of glial cells
CC (e.g. multiple sclerosis). They can be used to treat patients whose
CC nervous system has been damaged by e.g. trauma, surgery, stroke,
CC ischaemia, infection, metabolic disease, nutritional deficiency,
CC malignancy, or toxic agents. NRG3 can also be used to treat

CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangier's disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease of skeletal muscle of smooth muscle, such as muscular dystrophy or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

Query Match 100.0%; Score 277; DB 1; Length 720;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 47
|||||
DB 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 332
|||||

RESULT 4

ID W97621 standard; Protein; 360 AA.
AC W97621;
DT 10-MAY-1999 (first entry)
DE Human neuregulin related ligand NRG3 extracellular domain.
KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis.
KW Homo sapiens.
OS WO9902681-AL.
PN 21-JAN-1999.
PF 30-JUN-1998; U13411.
PR 24-JUL-1997; US-899437.
PR 09-JUL-1997; US-052019.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Zhang D;
DR WPI: 99-120882/10.
PT New isolated neuregulin related ligand-3 - used to develop products for treating nervous system disorders, e.g. stroke, ischaemia, infection, malignancy, Alzheimer's disease or Down's syndrome.
PS Claim 5(a); Page 69-70; 101pp; English.
CC This is the extracellular domain (ECD, aa1-360 of human neuregulin related ligand NRG3 (see also W97618); a novel member of the epidermal growth factor (EGF)-like family of protein ligands. NRG3 binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, activates ErbB4 receptor tyrosine phosphorylation. The invention provides human and murine polypeptides (see also W97617) that have at least 75% homology to the NRG3 ECD, as well as expression vectors, host cells and methods for the recombinant production of novel NRG3s. The NRG3 polypeptides and polynucleotides can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, ischaemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease,

CC and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangier's disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease of skeletal muscle of smooth muscle, such as muscular dystrophy or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

Query Match 100.0%; Score 277; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 47
|||||
DB 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 332
|||||

Query Match 100.0%; Score 277; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 47
|||||
DB 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYQGVRCDOFL 332
|||||

RESULT 5

ID W97622 standard; Protein; 47 AA.
AC W97622;
DT 10-MAY-1999 (first entry)
DE Human neuregulin related ligand NRG3 EGF-like domain.
KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis; epidermal growth factor; EGF;
KW immunoadhesin.
OS Homo sapiens.
PN WO9902681-AL.
PD 21-JAN-1999.
PF 30-JUN-1998; U13411.
PR 24-JUL-1997; US-899437.
PR 09-JUL-1997; US-052019.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Zhang D;
DR WPI: 99-120882/10.
PT New isolated neuregulin related ligand-3 - used to develop products for treating nervous system disorders, e.g. stroke, ischaemia, infection, malignancy, Alzheimer's disease or Down's syndrome.
PS Claim 30; Page 64; 101pp; English.
CC This is the epidermal growth factor (EGF)-like domain of human neuregulin related ligand NRG3 (see also W97618), a novel member of the EGF-like family of protein ligands that binds to the ErbB4 receptor and activates ErbB4 receptor tyrosine phosphorylation. The EGF-1 like domain of NRG3 is distinct from the EGF-like domains of NRG1 and NRG2. The invention provides human and murine polypeptides (see also W97617) that have at least 75% homology to the NRG3 EGF-like domain, as well as expression vectors, host cells and methods for the recombinant production of novel NRG3s. The NRG3 polypeptides and polynucleotides can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, ischaemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease,

CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
 CC disease, metachromatic leukodystrophy, Fabry's disease and
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
 CC smooth muscle, such as muscular dystrophy or diseases caused by
 CC skeletal or smooth muscle wasting. The products can also be used
 CC for detection, diagnosis, for the production of transgenic or
 CC knockout animals or for drug screening. A claimed immunoadhesin
 CC comprises the human NRG3 EGF-like domain fused to an immunoglobulin
 CC sequence. 47 AA;
 SQ

Query Match 100.0%; Score 277; DB 1; Length 47;
 Best Local Similarity 100.0%; Pred. No. 3e-22;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47
 |||||
 Db 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47

RESULT 6
 W97619 standard; Protein: 696 AA.
 ID W97619 (first entry)
 AC W97619
 DT 10-MAY-1999
 DE Human neuregulin related ligand NRG3 (splice variant).
 KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
 KW signal transduction; nervous system disorder; neurodegeneration;
 KW neuropathy; therapy; diagnosis; splice variant.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 1..360
 FT /note= "extracellular domain, specifically claimed
 FT in Claim 5(a)"
 FT Region 66..91
 FT /note= "hydrophobic region"
 FT Region 101..284
 FT /note= "mucin-like Ser/Thr-rich region, contains
 FT sites for O-linked glycosylation"
 FT Domain 285..354
 FT /note= "EGF-like domain"
 FT Domain 356..394
 FT /note= "transmembrane domain"
 FT W09902681-A1.
 PN 21-JAN-1999.
 PD 30-JUN-1998; U13411...
 PF 24-JUL-1997; US-899437...
 PR 09-JUL-1997; US-052019.
 PA (GETH) GENENTECH INC.
 PI Godowski P, Mark MR, Zhang D;
 DR WPI: 99-12082/10.
 DR N-PSDB: X06989.
 PT New isolated neuregulin related ligand-3 - used to develop products
 PT for treating nervous system disorders, e.g. stroke, ischaemia,
 PT infection, malignancy, Alzheimer's disease or Down's syndrome
 PS Example 1: Page 78-81; 101pp; English.
 CC This is the amino acid sequence of splice variant hNRG3B2 of human
 CC neuregulin related ligand NRG3, a novel member of the epidermal
 CC growth factor (EGF)-like family of protein ligands that binds to
 CC the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, and
 CC which activates ErbB4 receptor tyrosine phosphorylation. The
 CC sequence was deduced from the nucleotide sequence of a cDNA clone
 CC (see X06989) from a foetal brain library. hNRG3B2 lacks amino
 CC acids 329-552 of hNRG3B1 (see W97618) but retains the EGF-like
 CC domain and is expected to exhibit biological activity. The invention
 CC provides human and murine NRG3 polypeptides (see W97617), expression
 CC vectors, host cells and methods for the recombinant production of
 CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
 CC enhance the survival, proliferation or differentiation of cells
 CC having the ErbB4 receptor in vivo and in vitro. They can be used to
 CC prevent or treat damage to a nerve or damage to other NRG3-expressing
 CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In

CC particular, they can be used to treat diseases which involve neural
 CC cell growth such as demyelination, or damage or loss of glial cells
 CC (e.g. multiple sclerosis). They can be used to treat patients whose
 CC nervous system has been damaged by e.g. trauma, surgery, stroke,
 CC ischaemia, infection, metabolic disease, nutritional deficiency,
 CC malignancy, or toxic agents. NRG3 can also be used to treat
 CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou
 CC Gehrig's disease), Bell's palsy, conditions involving spinal
 CC muscular atrophy or paralysis, neurodegenerative disorders such as
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
 CC and Meniere's disease. They can also be used to treat neuropathies
 CC associated with systemic disease including post-polio syndrome,
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,
 CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
 CC disease, metachromatic leukodystrophy, Fabry's disease and
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
 CC smooth muscle, such as muscular dystrophy or diseases caused by
 CC skeletal or smooth muscle wasting. The products can also be used
 CC for detection, diagnosis, for the production of transgenic or
 CC knockout animals or for drug screening.
 SQ Sequence 696 AA;
 Query Match 100.0%; Score 277; DB 1; Length 696;
 Best Local Similarity 100.0%; Pred. No. 3.5e-21;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47
 |||||
 Db 286 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 332

RESULT 7
 Y05451 standard; Protein: 157 AA.
 ID Y05451 (first entry)
 AC Y05451
 DT 06-JUL-1999
 DE Human heregulin-like factor sequence.
 KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;
 KW neural system disorder; cancer.
 OS Homo sapiens.
 PN W09857989-A1.
 PD 23-DEC-1998.
 PF 16-JUN-1998; U12403.
 PR 17-JUN-1997; US-049942.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (GEOU) UNIV GEORGETOWN.
 PI Hijazi MM, King CR, Ruben SM, Young P;
 DR WPI: 99-095327/08.
 DR N-PSDB: X36423.
 PT New isolated heregulin-like factor - used to develop products for
 PT the diagnosis and treatment of disorders involving regulation of
 PT cell growth, particularly cancers
 PS Claim 17; Page 86-87; 118pp; English.
 CC This sequence is the human heregulin-like factor (HLF) of the
 CC invention. The HLF is involved in the regulation of cell growth.
 CC Detection of different levels of expression of the HLF gene can be used
 CC for the diagnosis of disorders, e.g. in the neural system. In
 CC particular, detection of different levels of HLF gene expression in cells
 CC or body fluid of an individual can be used for diagnosing cancer. The
 CC products can also be used in the treatment of disorders involving
 CC abnormal levels of HLF activity.
 SQ Sequence 157 AA;
 Query Match 100.0%; Score 277; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 9e-22;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 47
 |||||
 Db 31 HFPCRDLDLAYCLNDGECFVIETLTGSHKRCCKEGYGVRCDOFL 77

RESULT	9	
W05182	ID	W05182 standard; peptide; 52 AA.
AC	W05182;	
DT	04-JUN-1997	(first entry)
DE	Neu differentiation factor/heregulin-alpha/beta form EGF-like domain.	
ND	NDF; neu differentiation factor; heregulin; epidermal growth factor;	
KW	EGF; colon epithelial cell proliferation; Schwann cell; nerve;	
KW	damage; colitis; ulcer.	
OS	Synthetic.	
PN	W09631599-AI.	
PD	10-OCT-1996.	
PF	27-MAR-1996;	U04262.
PR	06-APR-1995;	US-417640.
PA	(AMGE-) AMGEN INC.	
PI	Cranahan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;	
DR	WPI: 96-465022/46.	
PT	Peptide(s) derived from neu differentiation factor/heregulin	
PT	Peptides - specifically from epidermal growth factor-like domain,	
PT	stimulate proliferation of colon epithelial cells and Schwann cells	
PS	Claim 1: Page 27: 37pp: English.	
CC	The peptides W05182-W05185 are based on neu differentiation factor	
CC	(NDF)/heregulin alpha and beta form EGF-like domains in various	
CC	combinations. The peptides maintain the survival and proliferation of	
CC	Schwann cells and cause proliferation, growth and differentiation of	
CC	colon epithelial cells. Accordingly, they are useful to treat (in vitro	
CC	or in vivo) a disease or disorder of the colon (e.g. colitis or an	
CC	ulcer) or of the nervous system (e.g. nerve damage caused by trauma).	
CC	Sequence	52 AA;
SQ		

Glial mitogenic polypeptide factors - used in diagnosis, prophylaxis or treatment of patho-physiological conditions of the nervous system
PT Disclosure; Fig 28b; 98pp; English.
PS
CC The sequences given in R28536-38 are encoded by cDNA sequences which have been derived from the putative bovine glial growth factor-II (GGF-II) cDNA by alternative splicing. The bovine GGF-II cDNA sequence was isolated by probing a bovine cDNA library with the sequences given in Q30583-616 which were designed by backtranslating amino acid sequences derived from purified GGF-II protein. Isolated open reading frames were amplified using polymerase chain reaction on pituitary RNA (see also Q30638-667). The isolated sequence was found to contain at least five exons and three alternative splice patterns could be used to produce the putative bovine GGF-II cDNA sequences 1, 2 and 3.
CC
CC Sequence 263 AA:
SO

```

Query Match      41.0%; Score 113.5; DB 1; Length 263;
Best Local Similarity 34.8%; pred. NO.47.4e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKD1AYCINDGECFVIETLTGSHKH-CRCKEGYQGVRCQD 45
      | : : : : | : | : | : | : : | : | : | : | : | : | : | : |
Db 194 HLVKCAEKETFCVNGCECFMWDLSPNSRYLCKPCQFGTGARCT 239

```

RESULT	11	
R29570		
ID	R29570 standard; Protein; 669 AA.	
AC	R29570;	
DC	20-APR-1993 (first entry)	
DE	Human heregulin-alpha deduced from clone lambda gt10 her16.	
DE	p185-HER2 ligand; her2 proto-oncogene; breast cancer; EGF receptor;	
KW	epidermal growth factor; tyrosine kinase-like glycoprotein;	
KW	HRG-alpha.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	domain	288..309
FT		/label= transmembrane
FT		/note= "probable"
FT	region	226..265
FT		/label= EGF motif
FT		

Db 2 HLVKCAEKETFCVNGGECFVWVRLDLSNPSRYLCKCQPGFTGARTC 47

RESULT 14

R46922 ID R46922 standard; Protein; 83 AA.
AC R46922; 28-JUL-1994 (first entry)
DE EGFL6.
KW Glial growth factor; GGF; heregulin; mitogenesis;
KW Schwann cell; tumour; central nervous system; erbB2 receptor;
KW antiproliferative; epidermal growth factor; EGF.
PN WO9403644-A.
PD 17-FEB-1994.
PF 10-AUG-1993; U07491.
PR 10-AUG-1992; US-927337.
PR 25-SEP-1992; US-951747.
PR 01-DEC-1992; US-984085.
PR 23-JAN-1993; US-011396.
PA (CAMP-) CAMBRIDGE NEUROSCIENCE INC.
PI Gwynne DI, Marchionni M, McBurney RN;
DR WPI; 94-065731/08.
DR N-PSDB; Q58328.
PT Glial growth factor DNA encoding numerous polypeptide factors
PT used for inhibiting cell proliferation - for treating carcinoma
PT and nervous disorders
PS Disclosure; Fig 44; 178pp: English.
CC The GGF coding segments include regions with EGF-like homology.
CC These EGF-like domains can be required for the activation of.
CC mitogenesis in the binding reaction between GGF ligands contg.
CC such domains and the erbB2 receptor. Pref. antiproliferative
CC factors are those which lack these EGF-like domains.
SQ Sequence 83 AA;

Query Match 41.0%; Score 113.5; DB 1; Length 83;
Best Local Similarity 34.8%; Pred. No. 2.6e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDO 45
Db 2 HLVKCAEKETFCVNGGECFVWVRLDLSNPSRYLCKCQPGFTGARTC 47

RESULT 15

R55659 ID R55659 standard; Protein; 63 AA.
AC R55659;
DT 28-JUL-1994 (first entry)
DE EGFL2.
KW Glial growth factor; GGF; heregulin; mitogenesis;
KW Schwann cell; tumour; central nervous system;
KW epidermal growth factor; EGF.
PN WO9400140-A.
PD 06-JAN-1994.
PF 29-JUN-1993; U06228.
PR 30-JUN-1992; US-907138.
PR 03-SEP-1992; US-940389.
PR 23-OCT-1992; US-965173.
PR 24-MAR-1993; US-036555.
PA (CAMP-) CAMBRIDGE NEUROSCIENCE.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Chen MS, Goodearl A, Hiles I, Marchioni M, Minghetti L;
PI Stroobant P, Waterfield M;
DR WPI; 94-025882/03.
DR N-PSDB; Q62843.
PT Glial mitogenic polypeptide factors - useful for stimulating
PT glial cell mitogenesis and treating glial cell tumours
PS Claim 53; Fig 39; 178pp: English.
CC EGFL1, EGFL2, EGFL3, EGFL4, EGFL5 and EGFL6 are used
CC for the stimulation of glial cell mitogenesis in vivo
CC and in vitro.

SQ Sequence 63 AA;

Query Match 41.0%; Score 113.5; DB 1; Length 63;
Best Local Similarity 34.8%; Pred. No. 2e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDO 45
Db 2 HLVKCAEKETFCVNGGECFVWVRLDLSNPSRYLCKCQPGFTGARTC 47

Search completed: October 13, 1999, 22:45:05
Job time: 215 sec

THIS PAGE BLANK (USPTO)

U.S. GOVERNMENT PRINTING OFFICE: 1964

RT "Cloning of cDNAs encoding xenopus neuroligin: expression in myotomal
muscle during embryo development."
RL Brain Res. Mol. Brain Res. 58:59-73(1998).

DR EMBL: AF076618; RAC28804.1; -
DR PFAM: PF000008; EGF; 1.
DR PFAM: PF00047; ig; 1.
KW Glycoprotein.

SQ SEQUENCE 677 AA; 75794 MW; 503C0COA CRC32;

Query Match 45.7%; Score 126.5; DB 13; Length 677;
Best Local Similarity 43.5%; Pred. No. 3.9e-09;
Matches 20; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45
DB 188 HLIKSDKEKTYCVNGGECYVLTGSHKH-CRCKEGYQGVRCDO 233

RESULT 3

Q12780 PRELIMINARY; PRT; 462 AA.

ID Q12780
AC Q12780
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE NEU DIFFERENTIATION FACTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94158863.

RA WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO Y.,
RA JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
RA MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
RA KOSKI R.A., YARDEN Y.,
RT "Structural and functional aspects of the multiplicity of Neu
RT differentiation factors."
RL Mol. Cell. Biol. 14:1909-1919(1994).
DR EMBL: U02326; AAA19951.1; -
DR PFAM: PF00008; EGF; 1.
DR PFAM: PF00047; ig; 1.
KW Glycoprotein.

SQ SEQUENCE 462 AA; 50878 MW; F16A34DC CRC32;

Query Match 41.0%; Score 113.5; DB 4; Length 462;
Best Local Similarity 34.8%; Pred. No. 1.5e-07;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45
DB 178 HLVKCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARTC 223

RESULT 4

Q12781 PRELIMINARY; PRT; 125 AA.

ID Q12781
AC Q12781
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE NEU DIFFERENTIATION FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94158863.

RA WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO Y.,
RA JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
RA MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,

RA KOSKI R.A., YARDEN Y.,
RT "Structural and functional aspects of the multiplicity of Neu
RT differentiation factors."
RL Mol. Cell. Biol. 14:1909-1919(1994).
DR EMBL: U02327; AAA19952.1; -
DR PFAM: PF00008; EGF; 1.
KW Glycoprotein.
FT NON_TER 1
SQ SEQUENCE 125 AA; 13137 MW; 4CB51B23 CRC32;

Query Match 41.0%; Score 113.5; DB 4; Length 125;
Best Local Similarity 34.8%; Pred. No. 4.2e-08;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45
DB 56 HLVKCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARTC 101

RESULT 5

Q02297 PRELIMINARY; PRT; 640 AA.

ID Q02297
AC Q02297
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE HERGULIN-ALPHA (HRG-ALPHA).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 2-22 AND 182-205.
RX MEDLINE; 92271253.
RA HOLMES W.E., SLIWKOWSKI M.X., AKITA R.W., HENZEL W.J., LEE J.,
RA PARK J.W., YANSURA D., ABADI N., RAAB H., LEWIS G.D., SHEPARD H.M.,
RA KUANG W.J., WOOD W.I., GOEDDEL D.V., VANDLEN R.L.,
RT "Identification of heregulin, a specific activator of p185erbB2."
RL Science 256:1205-1210(1992).
RN [2]

RP STRUCTURE BY NMR OF 175-241.
RA NAGATA K., KOHDA D., HATANAKA H., ICHIKAWA S., MATSUDA S.,
RA YAMAMOTO T., SUZUKI A., INAGAKI F.,
RL Submitted (JUL-1994) to the PDB data bank.
CC -!- SIMILARITY: BELONGS TO EGF FAMILY.

DR EMBL: M94165; AAA38638.1; -
DR PIR: A43273; A43273.
DR PDB: 1HRE; 15-OCT-94.
DR PDB: 1HRF; 15-OCT-94.
DR PROSITE: PS00022; EGF; 1.
DR PFAM: PF00008; EGF; 1.
DR PFAM: PF00047; ig; 1.
DR PFAM: PF00047; ig; 1.
KW Transmembrane; EGF-like domain.
FT DOMAIN 165 177 SER/THR-RICH.
FT TRANSMEM 243 265 POTENTIAL.
FT DISULFID 182 196
FT DISULFID 190 210
FT DISULFID 212 221

SQ SEQUENCE 640 AA; 70391 MW; 7A13DF1C CRC32;

Query Match 41.0%; Score 113.5; DB 4; Length 640;
Best Local Similarity 34.8%; Pred. No. 2e-07;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45
DB 178 HLVKCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARTC 223

RESULT 6

O35947 PRELIMINARY; PRT; 461 AA.
ID O35947

RESULT	8	
P43325		
ID	P43325	PRELIMINARY;
AC	P43325;	PRT; 461 AA.

RESULT		9	
P43326			
ID	P43326	PRELIMINARY;	PRT; 422 AA.
AC	P43326;		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-NOV-1998	(TREMBLrel. 08, Last annotation update)	
DE	NEU DIFFERENTIATION FACTOR NDF44.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
-RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE; 94158863.		
RA	WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO		
RA	JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,		
RA	MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,		
EA	KOSKI R.A., YARDEN Y.;		
RT	"Structural and functional aspects of the multiplicity of Neu		
RT	differentiation factors."		
RL	ML Cell. Biol. 14:1909-1919(1994).		
-RN	[2]		

DR EMBL; M92430;-; NOT_ANNOTATED_CDS.
DR PROSITE; PS00022; EGF; 1.
DR PFAM; PF00008; EGF; 1.
DR PFAM; PF00047; 15; 1.
KW Transmembrane; EGF-like domain.

FT	TRANSMEM	243	265	POTENTIAL.
FT	CONFLICT	90	90	K > N (IN REF. 2).
FT	CONFLICT	208	208	Y -> S (IN REF. 2).
SQ	SEQUENCE	422 AA;	46309 MW;	CCOE2197 CRC32;

```

Query Match      40.3%;   Score 111.5;   DB11;   Length 422;
Best Local Similarity 34.8%;   Pred. No. 2.5e-07;
Matches 16;   Conservative 13;   Mismatches 16;   Indels 1;

QY 1 HFKPCRDKLAYLNDGCEFTVETLTGSHKH-CRCKEGYQGVRCQD 45
      | | | | | | | | | | | | | | | | | | | | | |
Db 178 HLIIKAEKEKTFVYNGGSCFTVDLSNPSRYLCKQPGFTGARCT 223

```

RESULT	10
ID	Q12782
ID	Q12782 PRELIMINARY; PRT; 552 AA.
AC	Q12782;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-NOV-1998 (TrEMBLrel. 08, last annotation update)
DE	NEU DIFFERENTIATION FACTOR (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
NC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=PITUITARY;
RX	MEDLINE; 94158863.
RA	WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO
RA	JANSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
RA	MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
RA	KOSKI R.A., YARDEN Y.;
RT	"Structural and functional aspects of the multiplicity of Neu
RT	differentiation factors".
RL	Mol. Cell. Biol. 14:1909-1919(1994).
DR	EMBL; F00328; AAA19953.1; -
DR	PFAM; PF00008; EGF; 1.
FT	NON_TER 1
SQ	SEQUENCE 552 AA; 60845 MW; 3F7DD741 CRC32;

```

Query Match      37.7%  Score 104.5; DB 4; Length 552;
Best Local Similarity 31.2%  Pred. No. 2.8e-06;
Matches 15; Conservative 14; Mismatches 18; Indels 1;

QY 1 HFKPCRDKLAICLNDGSCFVIETLTGSHKH-CRCCKEGYQGVRCQDFL 47
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 85 HLVCASKEKTCFCVNGGCFMVKDLNSPRLVLCCKPNEFTGRCQNY 132
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT	11
Q12783	
ID	Q12783 PRELIMINARY; PRT; 288 AA.
AC	Q12783;
DT	01-NOV-1996 (TREMblrel. 01, Created)
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT	01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE	NEU DIFFERENTIATION FACTOR (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
[1]	
RX	SEQUENCE FROM N.A. MEDLINE: 94158863.
RA	WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LUO JANSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L., MENG S.Y., LO H.-S., HU S., CHANG D., YANG W., YANIGAHARA D., KOSKI R.A., YARDEN Y.;
RT	"Structural and functional aspects of the multiplicity of Neu differentiation factors."
RL	Mol. Cell. Biol. 14:1909-1919(1994).

DR	EMBL; U02329; AAA19954.1; -
DR	PFAM; PF00008; EGF; 1
FT	NON_TER 1 1
FT	NON_TER 288 288
SQ	SEQUENCE 288 AA; 31488 MW; 60DF34F9 CRC32;

```

Query Match      37.7%; Score 104.5; DB 4; Length 288;
Best Local Similarity 31.2%; Pred.No 1.5e-06;
Matches 15; Conservative 14; Mismatches 18; Indels 1
Qy 1 HFKPCRDLAYLNDGCECFVETLTGSHKI-CRCKEGYQVGRCDQFL 47
      + + + + + : : : : : : : : : : : : : : : : : : : : : :
Db 60 HLVAKAEKTEFTGVNGGCEFMVYDKLSNPSRYLKACPNFETGDRCONV 107
      + + + + + : : : : : : : : : : : : : : : : : : : : : :

```

RESULT	12
Q12784	
ID	PRELIMINARY;
AC	PRT; 175 AA.
Q12784;	
01-NOV-1996	(TRENBLrel. 01, Created)
DT	01-NOV-1996 (TRENBLrel. 01, last sequence update)
DT	01-NOV-1998 (TRENBLrel. 08, last annotation update)
DE	NEU DIFFERENTIATION FACTOR (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE; 94158863.
RA	WEN D., SUGGS S.V., KARUNAGARAN D., LIU N., CUPPLES R.L., LJO
RA	JANSSEN A.M., BEN-BARUCH N., TROLLINGER D.B., JACOBSEN V.L.,
RA	MENG S.Y., LU H.S., HU S., CHANG D., YANG W., YANIGAHARA D.,
RA	KOSKI R.A., YARDEN Y.;
RT	"Structural and functional aspects of the multiplicity of Neu
RT	differentiation factors.";
RT	Mol. Cell. Biol. 14:1909-1919(1994).
DR	EMBL; U02330; AAA19955.1; -.
DR	PFAM; PF00008; EGF; 1.
DR	NON_TER 1
SQ	SEQUENCE 175 AA; 18947 MW BE170E77 CRC32:

Query Match 37.7% ; Score 104.5; DB 4; Length 175;
Best Local Similarity 31.2%; Pred.No.9.2e-08;
Matches 15; Conservative 14; Mismatches 1

QY 1 HFKPCRDLAYCLNDGSCFVIETLTGTSHKI-CRCKEGYGVGRCDQFL 47

Db 112 HLVKCAEKTEFCVNGEGCFMVKDLSPNRYLCKCPNEFTGDRCNVY 159

RESULT	13
ID	Q15491
AC	PRELIMINARY; PRT; 296 AA.
DC	Q15491;
DT	01-NOV-1996 (TrenBLrel. 01, Created)
DT	01-NOV-1996 (TrenBLrel. 01, Last sequence update)
DT	01-NOV-1998 (TrenBLrel. 08, Last annotation update)
DE	SENSORY AND MOTOR NEURON-DERIVED FACTOR.
DE	SMDF.
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN STEM;
RX	MEDLINE; 95301541;
RA	HO W.H., ARMANINI M.P., NIJENS A., PHILLIPS H.S., OSHEROFF P
RT	"Sensory and motor neuron-derived factor. A novel heregulin v
RT	highly expressed in sensory and motor neurons.";
J. Biol. Chem.	270:14523-14532(1995).
EMBL:	L41827; AAC41764.1; -
DR	

CC -!- SIMILARITY: BELONGS TO EGF FAMILY.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 1999, 14:00:21 ; Search time 33.23 Seconds
(without alignments)
56.668 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFKPCRDLDLAYCLNDGECF.....SHKHCRCKEGYQGVRCDOFL 27

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR-60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	113.5	41.0	640	2	A43273	heregulin precursor
2	113.5	41.0	462	2	I38404	neu differentiation
3	113.5	41.0	125	2	I38405	neu differentiation
4	111.5	40.3	461	2	I61716	neu differentiation
5	111.5	40.3	639	2	I61719	neu differentiation
6	111.5	40.3	422	2	I61724	neu differentiation
7	105.5	38.1	125	2	S62676	heregulin isoform
8	104.5	37.7	422	2	S32357	glial growth facto
9	104.5	37.7	637	2	C43273	heregulin precursor
10	104.5	37.7	241	2	D43273	heregulin precursor
11	104.5	37.7	645	2	B43273	heregulin, splice
12	104.5	37.7	552	2	I38406	neu differentiation
13	104.5	37.7	288	2	I38407	neu differentiation
14	104.5	37.7	175	2	I38408	neu differentiation
15	104.5	37.7	296	2	A56943	sensory/motor neur
16	104.5	37.7	241	2	S32359	glial growth facto
17	102.5	37.0	636	2	I61718	neu differentiation
18	102.5	37.0	636	2	I61721	neu differentiation
19	102.5	37.0	562	2	I61722	neu differentiation
20	102.5	37.0	230	2	A56210	neu differentiation
21	102.5	37.0	304	2	I61720	neu differentiation
22	98.5	35.6	602	2	A45769	acetylcholine rece
23	90	32.5	850	2	JC5700	Erbb kinase activa
24	90	32.5	868	2	JC5701	Erbb kinase activa
25	90	32.5	860	2	JC5702	Erbb kinase activa
26	90	32.5	641	2	PC4417	Erbb kinase activa
27	85	30.7	80	1	EGVZSF	growth factor - ra
28	84.5	30.5	1220	2	A56136	jagged protein pre
29	83	30.0	162	2	S68401	epiregulin precurs
30	83	30.0	46	2	JT0747	epiregulin - rat
31	83	30.0	2531	2	A46019	gene Notch-1 prote
32	83	30.0	861	2	A48825	Notch homolog Motc
33	82.5	29.8	177	2	A37408	betacellulin precu
34	82	29.6	85	1	EGVZM1	growth factor - my
35	82	29.6	230	2	A44074	probable EGF-like
36	80.5	29.1	178	2	JCL467	betacellulin precu
37	79.5	28.7	159	1	WFR1	transforming growt
38	79.5	28.7	140	1	WMV29	growth factor - va
39	79.5	28.7	142	1	WMV23C	growth factor - va

40 79.5 28.7 159 2 S27195 transforming growth
41 79.5 28.7 159 2 I57497 transforming growth
42 79 28.5 2531 2 S18188 notch protein homo
43 78 28.2 1207 1 EGHU epidermal growth f
44 77.5 28.0 722 2 I48324 DELTA-like 1 - mou
45 76 27.4 1217 1 EGM5MG epidermal growth f

ALIGNMENTS

RESULT 1
A43273
heregulin precursor, splice form alpha - human
N:Alternate names: breast cancer cell differentiation factor p45; Neu differentiation
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C:Accession: A43273; A48498; A38155
R:Holmes, W.E.; Silwakowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253
A:Accession: A43273
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-640 <HOL>
A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26
A:Note: sequence extracted from NCBI backbone (NCBIP:103250)
R:Culicouscou, J.M.; Plozman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.
J. Biol. Chem. 268, 18407-18410, 1993
A:Title: Characterization of a breast cancer cell differentiation factor that specifi
A:Reference number: A48498; MUID:93366731
A:Accession: A48498
A:Molecule type: protein
A:Residues: 20-21, 'X', 23-24, 'XX', 27-28 <CUL>
R:Peles, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Ya
Cell 69, 205-216, 1992
A:Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that ind
A:Reference number: A38155; MUID:92208945
A:Accession: A38155
A:Molecule type: protein
A:Residues: 'X', 13-16, 'X', 18-20, 'RG', 23-24, 'GP', 27, 'E', 29, 'XP', 32-36 <PEL>
A:Note: sequence extracted from NCBI backbone (NCBIP:91347)
C:Genetics:
A:Gene: GDB: HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: EGF homology
C:Keywords: alternative splicing; glycoprotein
F:182-221/Domain: EGF homology <EGF>

Query Match 41.0%; Score 113.5; DB 2; Length 640;
Best Local Similarity 34.8%; Pred. No. 5.6e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
Qy 1 HFKPCRDLDLAYCLNDGECFVIEITLTSKHK-CRCKEGYQGVRCDO 45
Db 178 HLVRCAERKEFCVNGGECFMVKDLSNPSRYLCKRCQPGTGARCTE 223

RESULT 2
I38404
neu differentiation factor - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I38404
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation
A:Reference number: A56210; MUID:94158863
A:Accession: I38404
A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA
A:Residues: 1-462 <RES>
A:Cross-references: EMBL:U02326; NID:g408402; PID:g408403

Query Match 41.0%; Score 113.5; DB 2; Length 462;
Best Local Similarity 34.8%; Pred. No. 4.2e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCRDLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDQ 45
| | :
DBB 178 HLVKCAEKTEFCVNGGECFMVKDLSNPRLCKCPGFTGARCTE 223

RESULT 3
I38405
neu differentiation factor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I38405
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863
A:Accession: I38405
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-125 <RES>
A:Cross-references: EMBL:U02327; NID:g408404; PID:g408405

Query Match 41.0%; Score 113.5; DB 2; Length 125;
Best Local Similarity 34.8%; Pred. No. 1.3e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCRDLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDQ 45
| | :
DBB 56 HLVKCAEKTEFCVNGGECFMVKDLSNPRLCKCPGFTGARCTE 101

RESULT 4
I61716
neu differentiation factor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Aug-1998
C:Accession: I61716; I61717
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863
A:Accession: I61716
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-461 <RES>
A:Cross-references: EMBL:U02316; NID:g408382; PID:g408383
A:Accession: I61717
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-461 <RE2>
A:Cross-references: EMBL:U02317; NID:g408384; PID:g408385

Query Match 40.3%; Score 111.5; DB 2; Length 461;
Best Local Similarity 34.8%; Pred. No. 7.1e-06;
Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFKPCRDLAYCLNDGECFVIETLTGSHKH-CRCKEGYGVRCDQ 45
| | :
DBB 178 HLVKCAEKTEFCVNGGECFTVKDLSNPRLCKCPGFTGARCTE 223

RESULT 5

A:Title: Isolation and structural characterization of recombinant human neu differentiation
 A:Reference number: S62676; MUID:96139341

A:Accession: S62676

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-67-16;17-30;31-38;39-58;59-92;93-120;121-125 <HAR>

C:Keywords: proto-oncogene

Query Match 38.1%; Score 105.5; DB 2; Length 125;
 Best Local Similarity 32.6%; Pred. No. 1.1e-05;
 Matches 15; Conservative 14; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVITLTGSHKH-CRCCKEGYQGVRCDOFL 45

DB 75 HLKCAEKEKTCVNGGECFMVKDLSNPRLCKCPQFTGARCTE 120

RESULT 8

A:Title: glial growth factor - human

A:Accession: S32357

C:Species: Homo sapiens (man)

C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999

A:Accession: S32357

R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Hen-

les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M.

Nature 362, 312-318, 1993

A:Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the n

A:Reference number: S32357; MUID:93205115

A:Accession: S32357

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-422 <MAR>

C:Superfamily: EGF homology

F:363-402/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 422;
 Best Local Similarity 31.2%; Pred. No. 4.3e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVITLTGSHKH-CRCCKEGYQGVRCDOFL 47

DB 359 HLKCAEKEKTCVNGGECFMVKDLSNPRLCKCPNEFTGRCQNYV 406

RESULT 9

A:Title: heregulin precursor, splice form beta-2 - human

A:Accession: C43273

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Aug-1998

A:Accession: C43273

R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu-

Science 256, 1205-1210, 1992

A:Title: Identification of heregulin, a specific activator of p185(erbB2).

A:Reference number: A43273; MUID:92271253

A:Accession: C43273

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: mRNA

A:Residues: 1-637 <HOL>

C:Genetics:

A:Gene: GDB:HGL

A:Cross-references: GDB:132656; OMIM:142445

A:Map position: 8p22-8p11

C:Superfamily: EGF homology

C:Keywords: alternative splicing

F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 637;
 Best Local Similarity 31.2%; Pred. No. 6.2e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVITLTGSHKH-CRCCKEGYQGVRCDOFL 47
 DB 178 HLKCAEKEKTCVNGGECFMVKDLSNPRLCKCPNEFTGRCQNYV 225

RESULT 10

DA3273

A:Title: heregulin precursor, splice form beta-3 - human

N:Alternate names: glial growth factor HRG-beta-3; heregulin

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999

A:Accession: D43273; S32358

R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan

Science 256, 1205-1210, 1992

A:Title: Identification of heregulin, a specific activator of p185(erbB2).

A:Reference number: A43273; MUID:92271253

A:Accession: D43273

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-241 <HOL>

R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.;

les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield,

Nature 362, 312-318, 1993

A:Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in th

A:Reference number: S32357; MUID:93205115

A:Accession: S32358

A:Molecule type: mRNA

A:Residues: 1-241 <MAR>

C:Genetics:

A:Gene: GDB:HGL; GGF

A:Cross-references: GDB:132656; OMIM:142445

A:Map position: 8p22-8p11

C:Superfamily: EGF homology

C:Keywords: alternative splicing

F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 241;
 Best Local Similarity 31.2%; Pred. No. 2.7e-05;
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVITLTGSHKH-CRCCKEGYQGVRCDOFL 47

DB 178 HLKCAEKEKTCVNGGECFMVKDLSNPRLCKCPNEFTGRCQNYV 225

RESULT 11

BA3273

A:Title: heregulin, splice form beta 1 - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Aug-1998

A:Accession: B43273

R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan

Science 256, 1205-1210, 1992

A:Title: Identification of heregulin, a specific activator of p185(erbB2).

A:Reference number: A43273; MUID:92271253

A:Accession: B43273

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-645 <HOL>

C:Genetics:

A:Gene: GDB:HGL

A:Cross-references: GDB:132656; OMIM:142445

A:Map position: 8p22-8p11

C:Superfamily: EGF homology

C:Keywords: alternative splicing

F:182-221/Domain: EGF homology <EGF>

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 1999, 01:54:24 ; Search time 10.23 Seconds
(without alignments)
129.874 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFKPCDKDLAYCLNDGECE.....SHKHCRCKEGYQVRCQDFL 47

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	32.5	850	1	NTAK_HUMAN
2	90	32.5	868	1	NTAK_RAT
3	85	30.7	80	1	GRFA_SFVKA
4	83	30.0	2531	1	BTC1_MOUSE
5	82.5	29.8	177	1	BTC_MOUSE
6	82	29.6	85	1	GRFA_MXVL
7	82	29.6	230	1	SPIT_DROME
8	80.5	29.1	178	1	BTC_HUMAN
9	79.5	28.7	142	1	GRFA_VACCC
10	79.5	28.7	140	1	GRFA_VACCV
11	79.5	28.7	159	1	TGFA_MOUSE
12	79.5	28.7	159	1	TGFA_RAT
13	79	28.5	2531	1	NTC1_RAT
14	78	28.2	1207	1	EGF_HUMAN
15	77.5	28.0	722	1	DLI1_MOUSE
16	77.5	28.0	714	1	DLI1_RAT
17	76	27.4	2139	1	CRB_DROME
18	76	27.4	1217	1	EGF_MOUSE
19	76	27.4	484	1	LEM2_PIG
20	75	27.1	611	1	LEM2_CANFA
21	74.5	26.9	723	1	DLI1_HUMAN
22	74.5	26.9	603	1	FAI2_CAVPO
23	74.5	26.9	1429	1	LI12_CAEEL
24	74	26.7	2871	1	FBNI_BOVIN
25	74	26.7	2871	1	FBNI_HUMAN
26	74	26.7	2871	1	FBNI_MOUSE
27	74	26.7	610	1	LEM2_HUMAN
28	74	26.7	551	1	LEM2_RABIT
29	73	26.4	294	1	GRK_DROME
30	73	26.4	409	1	MFGM_PIG
31	73	26.4	2524	1	NOTC_XENLA
32	73	26.4	1964	1	NTCA_MOUSE
33	72.5	26.2	615	1	FAI2_HUMAN
34	72.5	26.2	208	1	HBGF_MOUSE
35	72.5	26.2	463	1	MFGM_MOUSE
36	72.5	26.2	121	1	TGFA_MACMU
37	72	26.0	1133	1	EGF_RAT
38	72	26.0	570	1	FBP3_STRPU
39	72	26.0	768	1	LEM3_MOUSE
40	71.5	25.8	4543	1	LRP1_CHICK
41	71.5	25.8	160	1	TGFA_HUMAN
42	71.5	25.8	50	1	TGFA_RABIT
43	71	25.6	53	1	EGF_PIG

ALIGNMENTS

RESULT 1	NTAK_HUMAN	STANDARD:	PRT:	850 AA.
ID	NTAK_HUMAN			
AC	014511;			
DT	15-DEC-1998 (REL. 37, CREATED)			
DT	15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	NTAK PROTEIN (NEURAL- AND THYMUS- DERIVED ACTIVATOR FOR ERBB KINASES).			
GN	NTAK.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NEUROBLASTOMA;			
RX	MEDLINE: 98006324			
RA	HIGASHIYAMA S., HORIKAWA M., YAMADA K., ICHINO N., NAKANO N.,			
RA	NAKAGAWA T., MIYAGAWA J., MATSUSHITA N., NAGATSU T., TANIGUCHI N.,			
RA	ISHIGURO H.;			
RT	"A novel brain-derived member of the epidermal growth factor family			
RT	that interacts with ErbB3 and ErbB4."			
RL	J. BIOCHEM. 122:675-680(1997).			
CC	-!- FUNCTION: LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE RECEPTORS.			
CC	-!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS			
CC	1 C2-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO EGF FAMILY. TO HEPARIN-BINDING PROTEINS.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AB005060; J1024296; -			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PFAM; PF00008; EGF; 1.			
DR	PFAM; PF00047; I9; 1.			
KW	TRANSMEMBRANE; EGF-LIKE DOMAIN; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN.			
FT	TRANSMEM 406 426			
FT	POTENTIAL.			
FT	IG-LIKE C2-TYPE DOMAIN.			
FT	DOMAIN 250 318			
FT	EGF-LIKE.			
FT	DOMAIN 341 382			
FT	BY SIMILARITY.			
FT	DISULFID 257 311			
FT	BY SIMILARITY.			
FT	DISULFID 345 359			
FT	BY SIMILARITY.			
FT	DISULFID 372 381			
FT	BY SIMILARITY.			
FT	DOMAIN 10 13			
FT	POLY-PRO.			
FT	DOMAIN 20 30			
FT	POLY-SER.			
FT	DOMAIN 33 47			
FT	POLY-SER.			
FT	DOMAIN 87 90			
FT	POLY-ALA.			
FT	DOMAIN 721 727			
FT	POLY-PRO.			
FT	CARBOHYD 52 52			
FT	POTENTIAL.			
FT	CARBOHYD 53 53			
FT	POTENTIAL.			
FT	CARBOHYD 147 147			
FT	POTENTIAL.			
FT	CARBOHYD 278 278			
FT	POTENTIAL.			
FT	CARBOHYD 346 346			
FT	POTENTIAL.			
SQ	SEQUENCE 850 AA; 91678 MW; DES08427 CRC32;			

Query Match 32.5%; Score 90; DB 1; Length 850;
Best Local Similarity 34.9%; Pred. No. 0.0016;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

Query Match
Best Local Similarity 32.5%; Score 90; DB 1; Length 868;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

Query 1 HFPCRDKDLAYCLNDGECFVETLTGSHKHCRCKEYGVQVRC 43

DB 357 HARKCNETAKSYCVNGGYVIEGI--NQLSCKCPNGFFGQC 397

RESULT 3

ID GREFA_SFVKA STANDARD; PRT; 80 AA.

AC F08441;

DT 01-AUG-1988 (REL. 08, CREATED)

DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)

DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)

DE GROWTH FACTOR.

OS VIRUS FIBROMA VIRUS (STRAIN KASZA) (SFV).

OC VIROSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;

OC LEPOKIPOXVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 87172751.

RA CHANG W., UPTON C., HU S.-L., PURCHIO A.F., MCFADDEN G.;

RT "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains a growth factor gene with sequence similarity to those encoding epidermal growth factor and transforming growth factor alpha";

RL MOL. CELL. BIOL. 7:535-540(1987).

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; M15921; G333601; -

PRT; A26723; EGVSZF

DR PROSITE; PS00022; EGF_1; 1

DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR PFAM; PF00008; EGF; 1.

KW EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.

FT DOMAIN 29 73 EGF-LIKE.

FT DISULFID 33 47 BY SIMILARITY.

FT DISULFID 41 61 BY SIMILARITY.

FT DISULFID 63 72 BY SIMILARITY.

FT CARBOHYD 44 44 POTENTIAL.

FT CARBOHYD 54 54 POTENTIAL.

SQ SEQUENCE 80 AA: 9210 MW; AB4DA12B CRC32;

Query Match
Best Local Similarity 30.7%; Score 85; DB 1; Length 80;
Matches 20; Conservative 6; Mismatches 17; Indels 8; Gaps 3;

Query 1 HFPCRDKDLAYCLNDGECFVIE---TLGSHKHCRCKEYGVQVRC 47

DB 29 HVKVCNHDIYCNLNGTCFTIALDNVSIT---PFCVCINYEGRSC-QFI 75

FT CARBOHYD 163 163 POTENTIAL.

FT CARBOHYD 294 294 POTENTIAL.

FT CARBOHYD 362 362 POTENTIAL.

FT VARSPLIC 388 388 C->G (IN ISOFORM NTAK-GAMMA).

FT VARSPLIC 389 868 MISSING (IN ISOFORM NTAK-GAMMA).

FT VARSPLIC 390 421 NGFGQRCLKLEPLRMLYPDPKOKHLGLFELKE --> VGVTG DRCOQFAVFNSK (IN ISOFORM NTAK-BETA).

FT VARSPLIC 414 421 MISSING (IN ISOFORM NTAK-ALPHA2A).

SQ SEQUENCE 868 AA: 93776 MW; C202303C CRC32;

Query Match
Best Local Similarity 34.9%; Pred. No. 0.0017;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

Query 1 HFPCRDKDLAYCLNDGECFVETLTGSHKHCRCKEYGVQVRC 43

DB 357 HARKCNETAKSYCVNGGYVIEGI--NQLSCKCPNGFFGQC 397

RESULT 3

ID GREFA_SFVKA STANDARD; PRT; 80 AA.

AC F08441;

DT 01-AUG-1988 (REL. 08, CREATED)

DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)

DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)

DE GROWTH FACTOR.

OS VIRUS FIBROMA VIRUS (STRAIN KASZA) (SFV).

OC VIROSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;

OC LEPOKIPOXVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 87172751.

RA CHANG W., UPTON C., HU S.-L., PURCHIO A.F., MCFADDEN G.;

RT "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains a growth factor gene with sequence similarity to those encoding epidermal growth factor and transforming growth factor alpha";

RL MOL. CELL. BIOL. 7:535-540(1987).

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; M15921; G333601; -

PRT; A26723; EGVSZF

DR PROSITE; PS00022; EGF_1; 1

DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR PFAM; PF00008; EGF; 1.

KW EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.

FT DOMAIN 29 73 EGF-LIKE.

FT DISULFID 33 47 BY SIMILARITY.

FT DISULFID 41 61 BY SIMILARITY.

FT DISULFID 63 72 BY SIMILARITY.

FT CARBOHYD 44 44 POTENTIAL.

FT CARBOHYD 54 54 POTENTIAL.

SQ SEQUENCE 80 AA: 9210 MW; AB4DA12B CRC32;

Query Match
Best Local Similarity 30.7%; Score 85; DB 1; Length 80;
Matches 20; Conservative 6; Mismatches 17; Indels 8; Gaps 3;

Query 1 HFPCRDKDLAYCLNDGECFVIE---TLGSHKHCRCKEYGVQVRC 47

DB 29 HVKVCNHDIYCNLNGTCFTIALDNVSIT---PFCVCINYEGRSC-QFI 75

FT CARBOHYD 163 163 POTENTIAL.

FT CARBOHYD 294 294 POTENTIAL.

FT CARBOHYD 362 362 POTENTIAL.

FT VARSPLIC 388 388 C->G (IN ISOFORM NTAK-GAMMA).

FT VARSPLIC 389 868 MISSING (IN ISOFORM NTAK-GAMMA).

FT VARSPLIC 390 421 NGFGQRCLKLEPLRMLYPDPKOKHLGLFELKE --> VGVTG DRCOQFAVFNSK (IN ISOFORM NTAK-BETA).

FT VARSPLIC 414 421 MISSING (IN ISOFORM NTAK-ALPHA2A).

SQ SEQUENCE 868 AA: 93776 MW; C202303C CRC32;

Query Match
Best Local Similarity 34.9%; Pred. No. 0.0017;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

Query 1 HFPCRDKDLAYCLNDGECFVETLTGSHKHCRCKEYGVQVRC 43

DB 357 HARKCNETAKSYCVNGGYVIEGI--NQLSCKCPNGFFGQC 397

RESULT 3

ID GREFA_SFVKA STANDARD; PRT; 80 AA.

AC F08441;

DT 01-AUG-1988 (REL. 08, CREATED)

DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)

DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)

DE GROWTH FACTOR.

OS VIRUS FIBROMA VIRUS (STRAIN KASZA) (SFV).

OC VIROSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;

OC LEPOKIPOXVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 87172751.

RA CHANG W., UPTON C., HU S.-L., PURCHIO A.F., MCFADDEN G.;

RT "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains a growth factor gene with sequence similarity to those encoding epidermal growth factor and transforming growth factor alpha";

RL MOL. CELL. BIOL. 7:535-540(1987).

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; M15921; G333601; -

PRT; A26723; EGVSZF

DR PROSITE; PS00022; EGF_1; 1

DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR PFAM; PF00008; EGF; 1.

KW EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.

FT DOMAIN 29 73 EGF-LIKE.

FT DISULFID 33 47 BY SIMILARITY.

FT DISULFID 41 61 BY SIMILARITY.

FT DISULFID 63 72 BY SIMILARITY.

FT CARBOHYD 44 44 POTENTIAL.

FT CARBOHYD 54 54 POTENTIAL.

SQ SEQUENCE 80 AA: 9210 MW; AB4DA12B CRC32;

Query Match
Best Local Similarity 30.7%; Score 85; DB 1; Length 80;
Matches 20; Conservative 6; Mismatches 17; Indels 8; Gaps 3;

Query 1 HFPCRDKDLAYCLNDGECFVIE---TLGSHKHCRCKEYGVQVRC 47

DB 29 HVKVCNHDIYCNLNGTCFTIALDNVSIT---PFCVCINYEGRSC-QFI 75

FT CARBOHYD 163 163 POTENTIAL.

FT CARBOHYD 294 294 POTENTIAL.

FT CARBOHYD 362 362 POTENTIAL.

FT VARSPLIC 388 388 C->G (IN ISOFORM NTAK-GAMMA).

FT VARSPLIC 389 868 MISSING (IN ISOFORM NTAK-GAMMA).

FT VARSPLIC 390 421 NGFGQRCLKLEPLRMLYPDPKOKHLGLFELKE --> VGVTG DRCOQFAVFNSK (IN ISOFORM NTAK-BETA).

FT VARSPLIC 414 421 MISSING (IN ISOFORM NTAK-ALPHA2A).

SQ SEQUENCE 868 AA: 93776 MW; C202303C CRC32;

Query Match
Best Local Similarity 34.9%; Pred. No. 0.0017;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

Query 1 HFPCRDKDLAYCLNDGECFVETLTGSHKHCRCKEYGVQVRC 43

DB 357 HARKCNETAKSYCVNGGYVIEGI--NQLSCKCPNGFFGQC 397

RESULT 3

ID GREFA_SFVKA STANDARD; PRT; 80 AA.

AC F08441;

DT 01-AUG-1988 (REL. 08, CREATED)

DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)

DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)

DE GROWTH FACTOR.

OS VIRUS FIBROMA VIRUS (STRAIN KASZA) (SFV).

OC VIROSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;

OC LEPOKIPOXVIRUS.

RN [1]</

```
RESULT 4
NTCL_MOUSE STANDARD; PRT; 2531 AA.
AC Q01705;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
[1]
SEQUENCE FROM N.A.
RN
RP TISSUE=EMBRYO;
RX MEDLINE; 93194170.
RA FRANCO DEL AMO F., GENDRON-MAGUIRE M., SWIATEK P.J., JENKINS N.A.,
RA COPELAND N.G., GRIDLEY T.;
RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT homolog of Drosophila Notch.";
RL GENOMICS 15:259-264(1993).
[2]
SEQUENCE OF 1551-2170 FROM N.A.
RN
RP TISSUE=EMBRYO;
RX MEDLINE; 93048835.
RA FRANCO DEL AMO F., SMITH D.E., SWIATEK P.J., GENDRON-MAGUIRE M.,
RA GREENSPAN R.J., MCMAHON A.P., GRIDLEY T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RT development.";
RL DEVELOPMENT 115:737-744(1992).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; Z11886; G288503;
CC MGD; MGI:97363; NOTCH1.
CC PROSITE; PS00010; ASX_HYDROXYL; 22.
CC PROSITE; PS00022; EGF_1; 34.
CC PROSITE; PS01186; EGF_2; 27.
CC PROSITE; PS01187; EGF_CA; 21.
CC PFAM; PF00008; EGF; 35.
CC PFAM; PF00023; ank; 6.
CC PFAM; PF00066; notch; 3.
CC HSP; P00740; LIXA.
CC DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;
CC TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 18
FT CHAIN 19 2531
FT DOMAIN 19 1725
FT TRANSMEM 1726 1746
FT DOMAIN 1747 2531
FT DOMAIN 24 1425
FT DOMAIN 1449 1462
FT DOMAIN 1445 1562
FT REPEAT 1445 1480
FT REPEAT 1481 1522
FT REPEAT 1523 1562
FT DOMAIN 1865 2075
FT REPEAT 1865 1910
FT REPEAT 1912 1942
FT REPEAT 1944 1975
FT CHAIN 19 2531
FT TRANSMEM 1726 1746
FT DOMAIN 1747 2531
FT DOMAIN 24 1425
FT DOMAIN 1449 1462
FT DOMAIN 1445 1562
FT REPEAT 1445 1480
FT REPEAT 1481 1522
FT REPEAT 1523 1562
FT DOMAIN 1865 2075
FT REPEAT 1865 1910
FT REPEAT 1912 1942
FT REPEAT 1944 1975
```

```
FT REPEAT 1978 2009
FT REPEAT 2011 2042
FT REPEAT 2044 2075
FT CARBOHYD 888 888
FT CARBOHYD 959 959
FT CARBOHYD 1179 1179
FT CARBOHYD 1241 1241
FT CARBOHYD 1489 1489
FT CARBOHYD 1587 1587
SQ SEQUENCE 2531 aa; 271312 MW; AD71189B CRC32;

Query Match 30.0%; Score 83; DB 1; Length 2531;
Best Local Similarity 41.7%; Pred. NO. 0.03;
Matches 15; Conservative 4; Mismatches 13; Indels 4; Gaps 1;

QY 9 DLAYCLNDCGCVVETLTGSHKHCKEGYQGVRC D 44
DB 1064 DSAPCKNGRCW----QTNTQYHCECRSGTWGNC D 1095

RESULT 5
BTC_MOUSE STANDARD; PRT; 177 AA.
ID BTC_MOUSE
AC Q05928;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE BETACELLULIN PRECURSOR (BTC).
GN BTC OR BCN.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 32-54; 64-71 AND 75-111.
RN
RP TISSUE=PANCREAS;
RX MEDLINE; 93206093.
RA SHING Y., CHRISTOFORI G., HANAHAN D., ONO Y., SASADA R.,
RA IGARASHI K., FOLKMAN J.;
RT "Betacellulin: a mitogen from pancreatic beta cell tumors.";
RL SCIENCE 259:1604-1607(1993).
CC -!- FUNCTION: POTENT MITOGEN FOR RETINAL PIGMENT EPITHELIAL CELLS
CC AND VASCULAR SMOOTH MUSCLE CELLS. THE EFFECTS OF BETACELLULIN
CC ARE PROBABLY MEDIATED BY THE EGF RECEPTOR AND OTHER RELATED
CC RECEPTORS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);
CC EXTRACELLULAR (MATURE FORM).
CC -!- TISSUE SPECIFICITY: FOUND IN SEVERAL MOUSE TISSUES INCLUDING
CC KIDNEY, UTERUS AND LIVER AS WELL AS IN BETA TUMOR CELL LINE AND
CC MCF-7 CELLS. IT IS NOT DETECTED IN THE BRAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: STRONG, TO HUMAN BETACELLULIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; L08394; G293853;
CC MGD; MGI:99439; BTC.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PFAM; PF00008; EGF; 1.
CC HSP; P01135; 1YUF.
CC GROWTH FACTOR; MITOGEN; GLYCOPROTEIN; EGF-LIKE DOMAIN; TRANSMEMBRANE;
CC SIGNAL.
FT SIGNAL 1 31
FT CHAIN 32 111
FT PROPEP 112 177
FT BETACELLULIN.
FT REMOVED IN MATURE FORM.
```

FT DOMAIN 32 118 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 119 139 POTENTIAL.
 FT DOMAIN 140 177 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 146 153 ARG/LYS-RICH (BASIC).
 FT CARBOHYD 34 34 POTENTIAL.
 FT CARBOHYD 42 42 POTENTIAL.
 FT CARBOHYD 52 52 POTENTIAL.
 FT DOMAIN 65 105 EGF-LIKE.
 FT DISULFID 69 82 BY SIMILARITY.
 FT DISULFID 77 93 BY SIMILARITY.
 FT DISULFID 95 104 BY SIMILARITY.
 SQ SEQUENCE 177 AA; 19664 MW; B2A6D489 CRC32;

Query Match 29.8%; Score 82.5; DB 1; Length 177;
 Best Local Similarity 39.1%; Pred. No. 0.0031;
 Matches 18; Conservative 9; Mismatches 14; Indels 5; Gaps 3;

OY 1 HKPCRDKDLAYCNDGEC-FVIELTGTSHKCRKGYGVRCDO 45
 DB 65 HFSRCQYKHYCIH-GRCFVDEQTPS---CICEKGYGARCER 106

RESULT 6
 ID GREX_MYXVL STANDARD; PRT; 85 AA.
 AC P08072;
 DT 01-AUG-1988 (REL. 08, CREATED)
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE GROWTH FACTOR (MGF).
 OS MYXOMA VIRUS (STRAIN LAUSANNE).
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC LEPOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 87141350.
 RA UPTON C., MACEN J.L., MCFADDEN G.;
 RT "Mapping and sequencing of a gene from myxoma virus that is related
 RT to those encoding epidermal growth factor and transforming growth
 RT factor alpha."
 RL J. VIROL. 61:1271-1275(1987).
 RN [2]
 RP SYNTHESIS, AND ACTIVITY.
 RX MEDLINE; 91182748.
 RA LIN Y.-Z., KE X.-H., TAM J.P.;
 RT "Synthesis and structure-activity study of myxoma virus growth
 RT factor."
 RL BIOCHEMISTRY 30:3310-3314(1991).
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M15806; G332300;
 DR PIR; A26131; EGVM21.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PFAM; PF00008; EGF; 1.
 KW EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.
 FT DOMAIN 33 77 EGF-LIKE.
 FT DISULFID 37 51
 FT DISULFID 45 65
 FT DISULFID 67 76
 FT CARBOHYD 48 48 POTENTIAL.
 FT CARBOHYD 58 58 POTENTIAL.
 SQ SEQUENCE 85 AA; 9629 MW; 195B4804 CRC32;

Query Match 29.6%; Score 82; DB 1; Length 85;
 Best Local Similarity 41.3%; Pred. No. 0.0019;
 Matches 19; Conservative 4; Mismatches 21; Indels 2; Gaps 2;

OY 3 KPCRDKDLAYCNDGECFVIELTGTSHKCRKGYGVRCDOQL 47
 DB 35 KLCNDYKNYCLNGCTCTVALNNVSLNPFACACHINYGVSRG-QFI 79

RESULT 7
 ID SPIT_DROME STANDARD; PRT; 230 AA.
 AC Q01083;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE PROTEIN SPITZ PRECURSOR.
 GN SPI.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
 OC DROSOPHILIDAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92354912.
 RA RUTLEDGE B.J., ZHANG K., BIER E., JAN Y.N., PERRIMON N.;
 RT "The Drosophila spitz gene encodes a putative EGF-like growth factor
 RT involved in dorsal-ventral axis formation and neurogenesis."
 RL GENES DEV. 6:1503-1517(1992).
 RN [2]
 RP FUNCTION.
 RX MEDLINE; 95134691.
 RA FREEMAN M.;
 RT "The spitz gene is required for photoreceptor determination in the
 RT Drosophila eye where it interacts with the EGF receptor."
 RL MECH. DEV. 48:25-33(1994).
 CC -!- FUNCTION: PROBABLE LIGAND THAT IS BOUND BY THE RECEPTOR TOP.
 CC INVOLVED IN A NUMBER OF UNRELATED DEVELOPMENTAL CHOICES, FOR
 CC EXAMPLE, DORSAL-VENTRAL AXIS FORMATION, GLIAL MIGRATION, SENSORY
 CC ORGAN DETERMINATION, AND MUSCLE DEVELOPMENT. IT IS REQUIRED FOR
 CC PHOTORECEPTOR DETERMINATION.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: EXPRESSED THROUGHOUT THE EMBRYO.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M95199; G158455;
 DR FLYBASE; FBgn0005672; spi.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PFAM; PF00008; EGF; 1.
 DR HSP; P01132; 1EPH.
 DR NEUROGENESIS; DEVELOPMENTAL PROTEIN; TRANSMEMBRANE; SIGNAL;
 KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 230
 FT DOMAIN 18 139 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 140 160 POTENTIAL.
 FT DOMAIN 161 230 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 61 66 POLY-THR.
 FT DOMAIN 74 118 EGF-LIKE.
 FT DOMAIN 182 185 POLY-ASP.
 FT DISULFID 78 93 BY SIMILARITY.
 FT DISULFID 87 106 BY SIMILARITY.
 FT DISULFID 108 117 BY SIMILARITY.


```

FT  DISULFID      77  93      BY SIMILARITY.
FT  DISULFID      95  104     BY SIMILARITY.
SQ  SEQUENCE      178 AA;  19746 MW;  788CE6B0 CRC32;  5;  Gaps  3;

Query Match      29.1%; Score 80.5; DB 1; Length 178;
Best Local Similarity 41.3%; Pred. No. 0.0054;
Matches 19; Conservative 5; Mismatches 17; Indels 5; Gaps 3;

QY  1 HFXPCRDKDLAYCLNDGEC-FVIELTGTSHKHCKCKEGYQGVRCQ 45
    ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  65 HFSRCPKQYKHCIK-GRCFRVAEQTPS---CVCDEGYIGARCE 106

RESULT  9
GREA_VACC      STANDARD;      PRT;      142 AA.
ID  GREA_VACC      AC  P20494;
DT  01-FEB-1991 (REL. 17, CREATED)
DT  01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT  01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE  GROWTH FACTOR.
GN  CLIR.
OS  VACCINIA VIRUS (STRAIN COPENHAGEN).
OC  VIRUSES; DSNIA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
ON  ORTHOPOXVIRUS.
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 91021027.
RA  GOEBEL S.-J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA  PAOLETTI E.;
RL  "The complete DNA sequence of vaccinia virus.";
RL  VIROLOGY 179:247-266(1990).
RN  [2]
RP  COMPLETE GENOME.
RA  GOEBEL S.-J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA  PAOLETTI E.;
RL  VIROLOGY 179:517-563(1990).
CC  1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; M35027; G335333;
CC  PIR; C42503; WMVZ3C
DR  PROSITE; PS00022; EGF_1; 1.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PFAM; PF00008; EGF; 1.
DR  HSSP; P01132; 1EPH.
KW  EGF-LIKE DOMAIN; GROWTH FACTOR; GLYCOPROTEIN.
FT  DOMAIN 41 81
FT  EGF-LIKE.
FT  DISULFID 45 58
FT  BY SIMILARITY.
FT  DISULFID 53 69
FT  BY SIMILARITY.
FT  DISULFID 71 80
FT  BY SIMILARITY.
FT  CARBOHYD 34 34
FT  POTENTIAL.
FT  CARBOHYD 95 95
FT  POTENTIAL.
SQ  SEQUENCE 142 AA; 15777 MW; 00AFEB3 CRC32;

Query Match      28.7%; Score 79.5; DB 1; Length 142;
Best Local Similarity 38.5%; Pred. No. 0.0058;
Matches 15; Conservative 6; Mismatches 15; Indels 3; Gaps 2;

QY  5 CRKDKLAYCNDGECFVIELTGTSHKHCKCKEGYQGVRC 43
    | : ||| : | : | : ||| : ||| : ||| :
Db  45 CGPEGDGCLF-GDCIHARDIDG--MYCRGSHGYTGIRC 80

```


RN SEQUENCE FROM N.A.
 RP MEDLINE: 85111164.
 RA LEE D.C., ROSE T.M., WEBB N.R., TODARO G.J.;
 RT "Cloning and sequence analysis of a cDNA for rat transforming growth
 RT factor alpha";
 RL NATURE 313:489-491(1985).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE: 90220597.
 RA BLASBAND A.J., ROGERS K.T., CHEN X., AZIZKHAN J.C., LEE D.C.;
 RT "Characterization of the rat transforming growth factor alpha gene
 RT and identification of promoter sequences";
 RL MOL. CELL. BIOL. 10:2111-2121(1990).
 [3]
 RN SEQUENCE OF 39-88.
 RP MEDLINE: 84121322.
 RA MARQUARDT H., HUNKAPILLER M.W., HOOD L.E., TODARO G.J.;
 RT "Rat transforming growth factor type 1: structure and relation to
 RT epidermal growth factor";
 RL SCIENCE 223:1079-1082(1984).
 [4]
 RN SEQUENCE OF 39-67.
 RP MEDLINE: 84061765.
 RA MASSAGUE J.;
 RT "Epidermal growth factor-like transforming growth factor. I.
 RT Isolation, chemical characterization, and potentiation by other
 RT transforming factors from feline sarcoma virus-transformed rat
 RT cells";
 RL J. BIOL. CHEM. 258:13606-13613(1983).
 CC -!- FUNCTION: TGF ALPHA IS A MITOGENIC POLYPEPTIDE THAT IS ABLE TO
 CC BIND TO THE EGF RECEPTOR AND TO ACT SYNERGISTICALLY WITH TGF BETA
 CC TO PROMOTE ANCHORAGE-INDEPENDENT CELL PROLIFERATION IN SOFT AGAR.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);
 CC EXTRACELLULAR (MATURE FORM).
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02004; G57338; -
 DR EMBL: M31075; G54526; -
 DR EMBL: M31076; G207282; -
 DR PIR: A01389; WERT1
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PFAM: PF00008; EGF; 1.
 DR HSP: P01135; 1YUF.
 KW GROWTH FACTOR; TRANSFORMING PROTEIN; MITOGEN; GLYCOPROTEIN;
 KW EGF-LIKE DOMAIN; TRANSFORMER; SIGNAL.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 38 REMOVED IN THE MATURE FORM.
 FT CHAIN 39 88 TRANSFORMING GROWTH FACTOR ALPHA.
 FT PROPEP 89 159 REMOVED IN MATURE FORM.
 FT DOMAIN 24 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 123 POTENTIAL.
 FT DOMAIN 124 159 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 42 82 EGF-LIKE.
 FT DISULFID 46 59 BY SIMILARITY.
 FT DISULFID 54 70 BY SIMILARITY.
 FT DISULFID 72 81 BY SIMILARITY.
 FT CARBOHYD 25 25 POTENTIAL.
 FT CONFLICT 28 28 S -> P (IN REF. 1).
 SQ SEQUENCE 159 AA; 16960 MW; F1206255 CRC32;

Query Match 28.7%; Score 79.5; DB 1; Length 159;
 Best Local Similarity 37.8%; Pred. No. 0.0065;

Matches 17; Conservative 5; Mismatches 18; Indels 5; Gaps 3;
 QY 1 HFKPCRDKDAYCLNDGEC-FVIEITLTGSHKCRCKEGYQGVRC 44
 DB 42 HFNKCPDSHTQICFH-GTCFLVQEE--EKPACVCHSGYGVRC 82
 RESULT 13
 NTC1 RAT
 ID NTC1 RAT STANDARD; PRT; 2531 AA.
 AC Q07008;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.
 GN NOTCH1.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SCHWANN CELL;
 RX MEDLINE: 92111383.
 RA WEINMASTER G., ROBERTS V.J., LEMKE G.;
 RT "A homolog of Drosophila Notch expressed during mammalian
 RT development";
 RL DEVELOPMENT 113:199-205(1991).
 CC -!- FUNCTION: REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER
 CC OF TISSUES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGHEST LEVELS OCCUR BETWEEN
 CC DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE
 CC ADULT.
 CC -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
 CC -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
 CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X57405; G57635; -
 DR PROSITE: PS00010; ASX_HYDROXYL; 22.
 DR PROSITE: PS00022; EGF_1; 35.
 DR PROSITE: PS01186; EGF_2; 26.
 DR PROSITE: PS01187; EGF_CA; 21.
 DR PFAM: PF00008; EGF; 35.
 DR PFAM: PF00023; ank; 6.
 DR PFAM: PF00066; notch; 3.
 DR HSP: P00740; 1IXA.
 KW DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;
 KW TRANSFORMER; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 2531 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
 FT DOMAIN 19 1723 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1724 1746 POTENTIAL.
 FT DOMAIN 1747 2531 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 20 58 EGF-LIKE 1.
 FT DOMAIN 59 99 EGF-LIKE 2.
 FT DOMAIN 102 139 EGF-LIKE 3.
 FT DOMAIN 140 176 EGF-LIKE 4.
 FT DOMAIN 178 216 EGF-LIKE 5.
 FT DOMAIN 218 255 EGF-LIKE 6.
 FT DOMAIN 257 293 EGF-LIKE 7.
 FT DOMAIN 295 333 EGF-LIKE 8.
 FT DOMAIN 335 371 EGF-LIKE 9.
 FT DOMAIN 372 410 EGF-LIKE 10.
 FT DOMAIN 412 450 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).

FT	DOMAIN	452	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	537	552	BY SIMILARITY.
FT	DOMAIN	490	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	554	563	BY SIMILARITY.
FT	DOMAIN	526	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	570	580	BY SIMILARITY.
FT	DOMAIN	564	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	575	589	BY SIMILARITY.
FT	DOMAIN	566	EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	591	600	BY SIMILARITY.
FT	DOMAIN	603	EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	602	618	BY SIMILARITY.
FT	DOMAIN	641	EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	617	627	BY SIMILARITY.
FT	DOMAIN	678	EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	629	638	BY SIMILARITY.
FT	DOMAIN	716	EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	645	655	BY SIMILARITY.
FT	DOMAIN	753	EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	650	664	BY SIMILARITY.
FT	DOMAIN	791	EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	666	675	BY SIMILARITY.
FT	DOMAIN	829	EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	682	693	BY SIMILARITY.
FT	DOMAIN	869	EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	687	702	BY SIMILARITY.
FT	DOMAIN	907	EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	704	713	BY SIMILARITY.
FT	DOMAIN	945	EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	720	730	BY SIMILARITY.
FT	DOMAIN	983	EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	725	739	BY SIMILARITY.
FT	DOMAIN	1021	EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	741	750	BY SIMILARITY.
FT	DOMAIN	1059	EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	757	768	BY SIMILARITY.
FT	DOMAIN	1097	EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	762	777	BY SIMILARITY.
FT	DOMAIN	1145	EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	779	788	BY SIMILARITY.
FT	DOMAIN	1183	EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	795	806	BY SIMILARITY.
FT	DOMAIN	1221	EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	800	815	BY SIMILARITY.
FT	DOMAIN	1265	EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	817	826	BY SIMILARITY.
FT	DOMAIN	1307	EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	833	844	BY SIMILARITY.
FT	DOMAIN	1346	EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	838	855	BY SIMILARITY.
FT	DOMAIN	1384	CYS-RICH.	FT	DISULFID	857	866	BY SIMILARITY.
FT	DOMAIN	1387	6 X ANK MOTIF REPEATS.	FT	DISULFID	873	884	BY SIMILARITY.
FT	DOMAIN	1426	ANK MOTIF 1.	FT	DISULFID	878	893	BY SIMILARITY.
FT	DOMAIN	1462	ANK MOTIF 2.	FT	DISULFID	895	904	BY SIMILARITY.
FT	DOMAIN	1499	ANK MOTIF 3.	FT	DISULFID	911	922	BY SIMILARITY.
FT	DOMAIN	1865	ANK MOTIF 4.	FT	DISULFID	916	931	BY SIMILARITY.
FT	DOMAIN	1944	ANK MOTIF 5.	FT	DISULFID	933	942	BY SIMILARITY.
FT	DOMAIN	1978	ANK MOTIF 6.	FT	DISULFID	937	998	BY SIMILARITY.
FT	DOMAIN	2011	BY SIMILARITY.	FT	DISULFID	992	1007	BY SIMILARITY.
FT	DOMAIN	2044	BY SIMILARITY.	FT	DISULFID	1009	1018	BY SIMILARITY.
FT	DOMAIN	24	BY SIMILARITY.	FT	DISULFID	1025	1036	BY SIMILARITY.
FT	DOMAIN	31	BY SIMILARITY.	FT	DISULFID	1030	1045	BY SIMILARITY.
FT	DOMAIN	48	BY SIMILARITY.	FT	DISULFID	1047	1056	BY SIMILARITY.
FT	DOMAIN	57	BY SIMILARITY.	FT	DISULFID	1063	1074	BY SIMILARITY.
FT	DOMAIN	63	BY SIMILARITY.	FT	DISULFID	1068	1083	BY SIMILARITY.
FT	DOMAIN	68	BY SIMILARITY.	FT	DISULFID	1085	1094	BY SIMILARITY.
FT	DOMAIN	87	BY SIMILARITY.	FT	DISULFID	1101	1122	BY SIMILARITY.
FT	DOMAIN	89	BY SIMILARITY.	FT	DISULFID	1116	1131	BY SIMILARITY.
FT	DOMAIN	106	BY SIMILARITY.	FT	DISULFID	1133	1142	BY SIMILARITY.
FT	DOMAIN	111	BY SIMILARITY.	FT	DISULFID	1149	1160	BY SIMILARITY.
FT	DOMAIN	129	BY SIMILARITY.	FT	DISULFID	1154	1169	BY SIMILARITY.
FT	DOMAIN	138	BY SIMILARITY.	FT	DISULFID	1171	1180	BY SIMILARITY.
FT	DOMAIN	144	BY SIMILARITY.	FT	DISULFID	1187	1198	BY SIMILARITY.
FT	DOMAIN	155	BY SIMILARITY.	FT	DISULFID	1192	1207	BY SIMILARITY.
FT	DOMAIN	164	BY SIMILARITY.	FT	DISULFID	1209	1218	BY SIMILARITY.
FT	DOMAIN	175	BY SIMILARITY.	FT	DISULFID	1225	1244	BY SIMILARITY.
FT	DOMAIN	182	BY SIMILARITY.	FT	DISULFID	1238	1253	BY SIMILARITY.
FT	DOMAIN	189	BY SIMILARITY.	FT	DISULFID	1255	1264	BY SIMILARITY.
FT	DOMAIN	206	BY SIMILARITY.	FT	DISULFID	1271	1284	BY SIMILARITY.
FT	DOMAIN	222	BY SIMILARITY.	FT	DISULFID	1276	1293	BY SIMILARITY.
FT	DOMAIN	227	BY SIMILARITY.	FT	DISULFID	1295	1304	BY SIMILARITY.
FT	DOMAIN	245	BY SIMILARITY.	FT	DISULFID	1311	1322	BY SIMILARITY.
FT	DOMAIN	254	BY SIMILARITY.	FT	DISULFID	1316	1334	BY SIMILARITY.
FT	DOMAIN	261	BY SIMILARITY.	FT	DISULFID	1336	1345	BY SIMILARITY.
FT	DOMAIN	266	BY SIMILARITY.	FT	DISULFID	1352	1363	BY SIMILARITY.
FT	DOMAIN	283	BY SIMILARITY.	FT	DISULFID	1357	1372	BY SIMILARITY.
FT	DOMAIN	299	BY SIMILARITY.	FT	DISULFID	1374	1383	BY SIMILARITY.
FT	DOMAIN	306	BY SIMILARITY.	FT	DISULFID	1391	1403	BY SIMILARITY.
FT	DOMAIN	323	BY SIMILARITY.					
FT	DOMAIN	332	BY SIMILARITY.					
FT	DOMAIN	339	BY SIMILARITY.					
FT	DOMAIN	344	BY SIMILARITY.					
FT	DOMAIN	359	BY SIMILARITY.					
FT	DOMAIN	361	BY SIMILARITY.					
FT	DOMAIN	370	BY SIMILARITY.					
FT	DOMAIN	376	BY SIMILARITY.					
FT	DOMAIN	381	BY SIMILARITY.					
FT	DOMAIN	398	BY SIMILARITY.					
FT	DOMAIN	400	BY SIMILARITY.					
FT	DOMAIN	416	BY SIMILARITY.					
FT	DOMAIN	423	BY SIMILARITY.					
FT	DOMAIN	440	BY SIMILARITY.					
FT	DOMAIN	456	BY SIMILARITY.					
FT	DOMAIN	461	BY SIMILARITY.					
FT	DOMAIN	478	BY SIMILARITY.					
FT	DOMAIN	487	BY SIMILARITY.					
FT	DOMAIN	494	BY SIMILARITY.					
FT	DOMAIN	499	BY SIMILARITY.					
FT	DOMAIN	516	BY SIMILARITY.					
FT	DOMAIN	525	BY SIMILARITY.					
FT	DOMAIN	532	BY SIMILARITY.					

Query Match 28.5%; Score 79; DB 1; Length 2531;

Best Local Similarity 38.9%; Pred. No. 0.089;

Matches 14; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY 9 DLAYCLNDECEFEVIELTGTSHKHCRCKEGYQGVRC 44

DB 1064 DSAPCKNGKCNW---QNTQYHCECRSGWTGFNC 1095

FT	PEPTIDE	771	1023		EPIDERMAL GROWTH FACTOR (UROGASTRONE).
FT	DOMAIN	314	335		EGF-LIKE 1.
FT	DOMAIN	336	396		EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	397	437		EGF-LIKE 3.
FT	DOMAIN	435	477		EGF-LIKE 4.
FT	DOMAIN	741	781		EGF-LIKE 5.
FT	DOMAIN	831	869		EGF-LIKE 6.
FT	DOMAIN	870	911		EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	912	952		EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	972	1013		EGF-LIKE 9.
FT	DISULFID	318	330		BY SIMILARITY.
FT	DISULFID	325	339		BY SIMILARITY.
FT	DISULFID	341	354		BY SIMILARITY.
FT	DISULFID	360	371		BY SIMILARITY.
FT	DISULFID	367	380		BY SIMILARITY.
FT	DISULFID	382	395		BY SIMILARITY.
FT	DISULFID	401	412		BY SIMILARITY.
FT	DISULFID	408	421		BY SIMILARITY.
FT	DISULFID	423	436		BY SIMILARITY.
FT	DISULFID	439	451		BY SIMILARITY.
FT	DISULFID	447	461		BY SIMILARITY.
FT	DISULFID	463	476		BY SIMILARITY.
FT	DISULFID	745	756		BY SIMILARITY.
FT	DISULFID	752	765		BY SIMILARITY.
FT	DISULFID	767	780		BY SIMILARITY.
FT	DISULFID	835	846		BY SIMILARITY.
FT	DISULFID	840	855		BY SIMILARITY.
FT	DISULFID	857	868		BY SIMILARITY.
FT	DISULFID	874	888		BY SIMILARITY.
FT	DISULFID	881	897		BY SIMILARITY.
FT	DISULFID	899	910		BY SIMILARITY.
FT	DISULFID	916	929		BY SIMILARITY.
FT	DISULFID	923	938		BY SIMILARITY.
FT	DISULFID	940	951		BY SIMILARITY.
FT	DISULFID	976	990		BY SIMILARITY.
FT	DISULFID	984	1001		POTENTIAL.
FT	DISULFID	1003	1012		POTENTIAL.
FT	CARBOHYD	38	38		POTENTIAL.
FT	CARBOHYD	104	104		POTENTIAL.
FT	CARBOHYD	117	117		POTENTIAL.
FT	CARBOHYD	148	148		POTENTIAL.
FT	CARBOHYD	324	324		POTENTIAL.
FT	CARBOHYD	404	404		POTENTIAL.
FT	CARBOHYD	596	596		POTENTIAL.
FT	CARBOHYD	815	815		POTENTIAL.
FT	CARBOHYD	926	926		POTENTIAL.
FT	VARIANT	708	708		I -> M.
FSQ	SEQUENCE	1207 AA;	133945 MW;	B762FD23 CRC32;	
 Query Match 28.2%; Score 78; DB 1; Length 1207;					
Best Local Similarity 50.0%; Pred. No. 0.06;					
Matches 16; Conservative 1; Mismatches 13; Indels 2; Gaps 1;					
QY	12 YCLNDGCEPVIEITLGTGSHKHCKECGYGVRC	43			
	:				
Db	983 YCLHDGVCVMEIAL--DKYACNCVGIIGERC	1012			
 RESULT 15					
DLL1_MOUSE STANDARD; PRT; 722 AA.					
ID	DLL1_MOUSE				
AC	Q61483;				
DT	01-NOV-1997 (REL. 35, CREATED)				
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)				
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)				
DE	DELTA-LIKE PROTEIN 1 PRECURSOR (DELTA).				
GN	DLL1.				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	RODENTIA; SCIURCGNATHI; MURIDAE; MURINAE; MUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				

INTERFERENCE

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 1999, 00:32:25 ; Search time 10.18 Seconds
(without alignments)
52.765 Million cell updates/sec

Title: US-09-107-979-4
Perfect score: 277
Sequence: 1 HFKPCRDKLAYCLNDGEF.....SHKHCRCKEGYQGVRCDFL 47

Scoring table: BLOSUM62

Searched: 119832 seqs, 11428610 residues

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/PCTUS9_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116.5	42.1	52	1	US-08-417-640A-1
2	116.5	42.1	52	1	US-08-760-815-1
3	116.5	42.1	52	2	US-08-761-038-1
4	113.5	41.0	669	1	US-07-847-743B-8
5	113.5	41.0	66	1	US-07-847-743B-10
6	113.5	41.0	669	1	US-07-847-743B-13
7	113.5	41.0	95	1	US-07-847-743B-14
8	113.5	41.0	625	1	US-07-847-743B-26
9	113.5	41.0	54	1	US-08-179-481-111
10	113.5	41.0	669	1	US-08-456-201-8
11	113.5	41.0	66	1	US-08-456-201-10
12	113.5	41.0	669	1	US-08-456-201-13
13	113.5	41.0	95	1	US-08-456-201-14
14	113.5	41.0	625	1	US-08-456-201-26
15	113.5	41.0	52	1	US-08-417-640A-3
16	113.5	41.0	52	1	US-08-760-815-3
17	113.5	41.0	669	2	US-08-330-161-11
18	113.5	41.0	95	2	US-08-330-161-12
19	113.5	41.0	669	2	US-08-456-241-8
20	113.5	41.0	66	2	US-08-456-241-10
21	113.5	41.0	669	2	US-08-456-241-13
22	113.5	41.0	95	2	US-08-456-241-14
23	113.5	41.0	625	2	US-08-456-241-26
24	113.5	41.0	669	2	US-08-440-401-11
25	113.5	41.0	95	2	US-08-440-401-12
26	113.5	41.0	669	2	US-08-419-878B-11
27	113.5	41.0	95	2	US-08-419-878B-12
28	113.5	41.0	52	2	US-08-761-038-3
29	113.5	41.0	669	3	PCT-US92-04295A-8
30	113.5	41.0	66	3	PCT-US92-04295A-10
31	113.5	41.0	669	3	PCT-US92-04295A-13
32	113.5	41.0	95	3	PCT-US92-04295A-14
33	113.5	41.0	625	3	PCT-US92-04295A-26
34	111.5	40.3	54	1	US-08-179-481-110
35	108.5	39.2	50	1	US-08-036-555B-152
36	108.5	39.2	50	1	US-08-469-569-152
37	108.5	39.2	45	1	US-08-168-091A-39
38	108.5	39.2	50	1	US-08-249-322A-152
39	108.5	39.2	50	2	US-08-469-526A-152

ALIGNMENTS

```
RESULT 1
US-08-417-640A-1
; Sequence 1, Application US/08417640A
; Patent No. 5670342
; GENERAL INFORMATION:
; APPLICANT: Carnahan, Josette F.
; APPLICANT: Hara, Shinichi
; APPLICANT: Lu, Hsieng S.
; APPLICANT: Mayer, John P.
; APPLICANT: Yoshinaga, Steven K.
; TITLE OF INVENTION: NDF Peptides
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,640A
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REFERENCE/DOCKET NUMBER: A-310
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-417-640A-1

Query Match 42.1%; Score 116.5; DB 1; Length 52;
Best Local Similarity 33.3%; Pred. No. 9.8e-08;
Matches 16; Conservative 15; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFKPCRDKLAYCLNDGEFVIETLTGSHKH-CRCKEGYQGVRCDFL 47
Db 2 HLYKCAKEKTFVNGCEFWKDLNPNRYLCKQCPGFTGRCQNYV 49.

RESULT 2
US-08-760-815-1
; Sequence 1, Application US/08760815
; Patent No. 5686415
; GENERAL INFORMATION:
; APPLICANT: Carnahan, Josette F.
; APPLICANT: Hara, Shinichi
; APPLICANT: Lu, Hsieng S.
; APPLICANT: Mayer, John P.
; APPLICANT: Yoshinaga, Steven K.
; TITLE OF INVENTION: NDF Peptides
; NUMBER OF SEQUENCES: 6
```

Sequence 152, App
Sequence 152, App
Sequence 148, App
Sequence 152, App
Sequence 40, Appl
Sequence 5, Appl1

```

; NAME: Mazza, Richard J.
; REFERENCE/DOCKET NUMBER: A-310
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-761-038-1

Query Match          42.1%; Score 116.5; DB 2; Length 52;
Best Local Similarity 33.3%; Pred.No. 9.8e-08;
Matches 16; Conservative 15; Mismatches 16; Indels 1

QY      1 HFKPCRDKLAYCLNDGECFVETLTGSHKH-CRCKEGYQGVRCDQFL 47
        | :|: |:|: ||||::|: :: |:|: |:|: ||: ::
DB       2 HLVKCAEKETCVNGEGCFMWKDLNPNRYLCKCPQFTGARCONYV 49

RESULT      4
US-07-847-743B-8
Sequence 8, Application US/07847743B
Patent No. 5367060
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Structure, Production and Use of
TITLE OF INVENTION: Heregulin
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSEE:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk.
```

```

SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/847,743B
FILING DATE: 19920306
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
FILING DATE: 24-May-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
FILING DATE: 25-Sep-1991
APPLICATION DATA:
APPLICATION NUMBER: 07/790801
FILING DATE: 08-NO. 5367060-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 712P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-847-743B-8

```


Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCRDLDAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 222 HLVKCAEKETFCVNGGECFMVKDLSNPSRYLCKCPQGTGARTC 267

RESULT 5

US-07-847-743B-10
; Sequence 10, Application US/07847743B
; Patent No. 5367060

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Structure, Production and Use of

; TITLE OF INVENTION: Heregulin

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/847,743B

; FILING DATE: 19920306

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/705256

; FILING DATE: 24-May-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/765212

; FILING DATE: 25-Sep-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/790801

; FILING DATE: 08-No. 5367060-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Hensley, Max D.

; REGISTRATION NUMBER: 27,043

; REFERENCE/DOCKET NUMBER: 712P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/266-1489

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 66 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

US-07-847-743B-10

Query Match 41.0%; Score 113.5; DB 1; Length 66;

Best Local Similarity 34.8%; Pred. No. 2.8e-07;

Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCRDLDAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45

| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 2 HLVKCAEKETFCVNGGECFMVKDLSNPSRYLCKCPQGTGARTC 47

RESULT 6

US-07-847-743B-13
; Sequence 13, Application US/07847743B
; Patent No. 5367060

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; TITLE OF INVENTION: Structure, Production and Use of

; TITLE OF INVENTION: Heregulin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/847,743B

; FILING DATE: 19920306

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/705256

; FILING DATE: 24-May-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/765212

; FILING DATE: 25-Sep-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/790801

; FILING DATE: 08-No. 5367060-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Hensley, Max D.

; REGISTRATION NUMBER: 27,043

; REFERENCE/DOCKET NUMBER: 712P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/266-1489

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 669 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

US-07-847-743B-13

Query Match 41.0%; Score 113.5; DB 1; Length 669;

Best Local Similarity 34.8%; Pred. No. 3.1e-06;

Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCRDLDAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDO 45

| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 222 HLVKCAEKETFCVNGGECFMVKDLSNPSRYLCKCPQGTGARTC 267

RESULT 7

US-07-847-743B-14

; Sequence 14, Application US/07847743B

; Patent No. 5367060

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; TITLE OF INVENTION: Structure, Production and Use of

; TITLE OF INVENTION: Heregulin

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

Query Match 41.0%; Score 113.5; DB 1; Length 54;
Best Local Similarity 34.8%; Pred. No. 2.3e-07;

Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIELTGSXKH-CRCKEGYGVRCDQ 45
| | | | | : | | | | | : | | : | | : | | :
Db 4 HLKCAEKETFCVNGECFVMDLSPNRYLCKCPGFTGARTC 49

RESULT 10

US-08-456-201-8
; Sequence 8, Application US/08456201
; Patent No. 5641869

GENERAL INFORMATION:

; APPLICANT: Vandlen, Richard L.,
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,201
; FILING DATE:

; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,145

FILING DATE:

; APPLICATION NUMBER: 07/880,917
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; PRIOR APPLICATION DATA: 07/765212
; APPLICATION NUMBER: 07/790801
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.

REGISTRATION NUMBER: 27,034

REFERENCE/DOCKET NUMBER: 712P4

TELEPHONE: 415/266-1489

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 669 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-456-201-8

Query Match 41.0%; Score 113.5; DB 1; Length 669;
Best Local Similarity 34.8%; Pred. No. 3.1e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIELTGSXKH-CRCKEGYGVRCDQ 45
| | | | | : | | | | | : | | : | | : | | :
Db 222 HLKCAEKETFCVNGECFVMDLSPNRYLCKCPGFTGARTC 267

RESULT 11

US-08-456-201-10

; Sequence 10, Application US/08456201

; Patent No. 5641869

GENERAL INFORMATION:

; APPLICANT: Vandlen, Richard L.,
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,201
; FILING DATE:

CLASSIFICATION: 530

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,145

FILING DATE:

; APPLICATION NUMBER: 07/880,917

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256

; PRIOR APPLICATION DATA: 07/765212

; APPLICATION NUMBER: 07/790801

; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.

; REGISTRATION NUMBER: 27,034

; REFERENCE/DOCKET NUMBER: 712P4

; TELEPHONE: 415/266-1489

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 66 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-456-201-10

Query Match 41.0%; Score 113.5; DB 1; Length 66;
Best Local Similarity 34.8%; Pred. No. 2.8e-07;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGECFVIELTGSXKH-CRCKEGYGVRCDQ 45
| | | | | : | | | | | : | | : | | : | | :
Db 2 HLKCAEKETFCVNGECFVMDLSPNRYLCKCPGFTGARTC 47

RESULT 12

US-08-456-201-13

; Sequence 13, Application US/08456201

; Patent No. 5641869

GENERAL INFORMATION:

; APPLICANT: Vandlen, Richard L.,

; APPLICANT: Holmes, William E.

; TITLE OF INVENTION: Structure, Production and Use of

; TITLE OF INVENTION: Heregulin 2 Ligands

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 625 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-456-201-26

Query Match 41.0%; Score 113.5; DB 1; Length 625;
Best Local Similarity 34.8%; Pred. NO. 2.9e-06;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

1 HFKPCRDKLAYCLNDGCECFVETLTGSHKH-CRCKEGYQGVRC DQ 45
 178 HLVKCAEKETFCVNGGCECFMVKDLSNPSRYLCKCPGFTGARCTE 223

RESULT	15
--------	----

US-08-417-640A-3
; Sequence 3, Application US/08417640A
; Patent No. 5670342

GENERAL INFORMATION:

APPLICANT: Carnahan, Josette F.
APPLICANT: Hara, Shinichi
APPLICANT: Lu, Hsieng S.
APPLICANT: Mayer, John P.
APPLICANT: Yoshinaga, Steven K.
TITLE OF INVENTION: NDF Peptides
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1940 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA

ZIP: 91320
COMPUTER READABLE FORM:

```

/
/ COMPILER: READERABLE FORM.
/
/ MEDIUM TYPE: Floppy disk
/
/ COMPUTER: IBM PC compatible
/
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/
/ CURRENT APPLICATION DATA:
/
/ APPLICATION NUMBER: US/08/417,640A
/
/

```

FILING DATE:

```

/ CLASSIFICATION: 436
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mazza, Richard J.
/ REFERENCE/DOCKET NUMBER: A-310
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 52 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-417-640A-3

```

Query Match 41.0%; Score 113.5; DB 1; Length 52;
Best Local Similarity 34.8%; Pred. NO. 2.2e-07;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

Qy 1 HFKPCRDKLAYCLNDGECFVETLTGSHKH-CRCKEGYQGVRCQ 45
 | | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 2 HLVKAEKEKTFCVNGGECFMVKDLNSPSRYLCKCPGGTGARCTE 47

Search completed: October 14, 1999, 00:43:10
Job time: 645 sec

THIS PAGE BLANK (USPTO)